Tracking pandemic virus threats in Thailand's bats p. 234

A subset of killer T cells help restrain human immune responses pp. 243 & 265

Aerosols can focus incident light internally p. 293

Science

$15
15 APRIL 2022
science.org

AAAS

PRINTED GLASS

Microscale components made using axial lithography p. 308

COVID-19 INNOVATORS WIN ROSS PRIZE

The Feinstein Institutes and Molecular Medicine awarded the **Ross Prize in Molecular Medicine** to University of Pennsylvania scientists for their contributions to developing COVID-19 mRNA-based vaccines.

Congratulations to:

KATALIN KARIKÓ

and

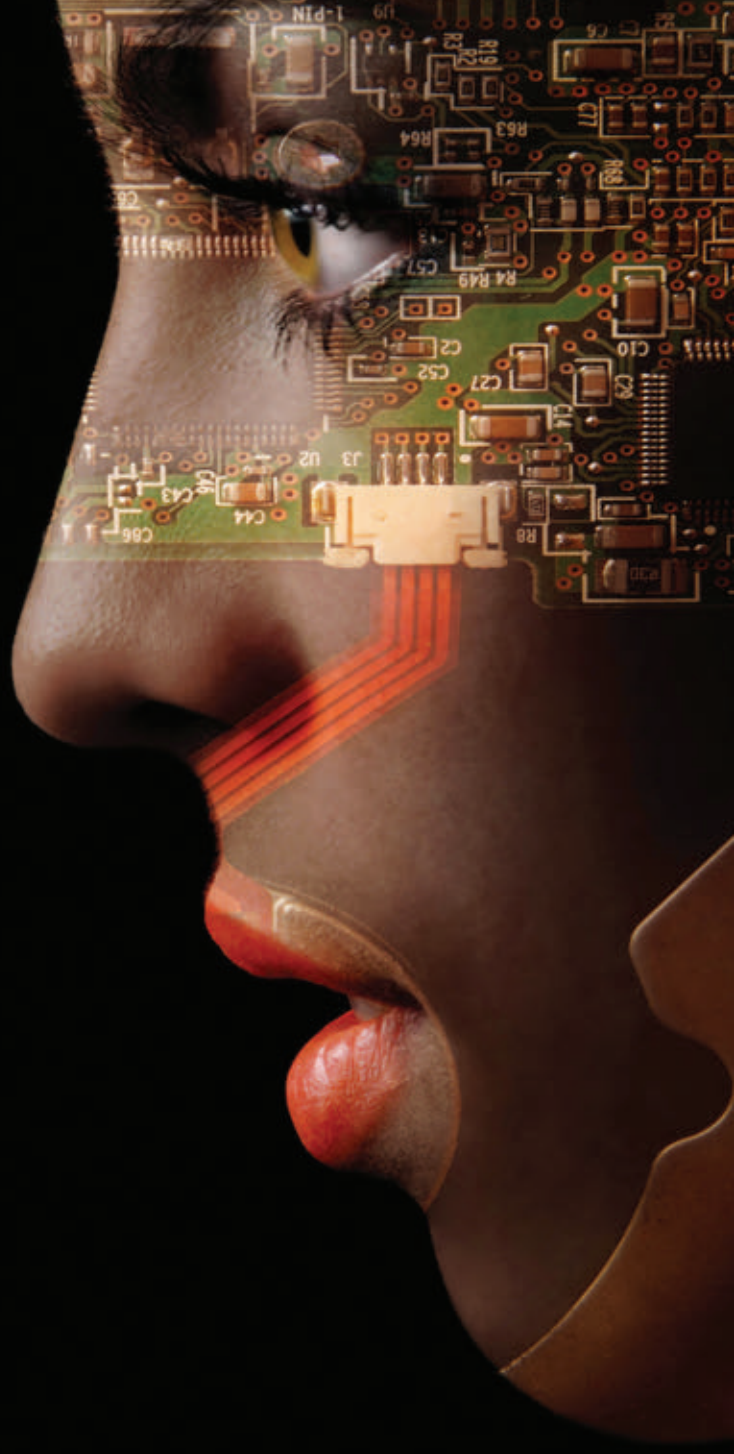
DREW WEISSMAN

We look forward to presenting the award during a research symposium in NYC on June 7.

To attend the symposium, or to submit a 2023 Ross Prize nominee, visit Northwell.edu/RossPrize

Feinstein Institutes for Medical Research
Northwell Health®

Share **Your Robotics Research** with the **World**.



Shaping the future of robotics with high impact research!

As a multidisciplinary online-only journal, *Science Robotics* publishes original, peer-reviewed, research articles that advance the field of robotics. The journal provides a central forum for communication of new ideas, general principles, and original developments in research and applications of robotics for all environments.

Submit your research today. Learn more at: **science.org/journal/scirobotics**

Twitter: @SciRobotics
Facebook: @ScienceRobotics

Science Robotics
AAAS

The lush lobby of CST headquarters in Danvers, Massachusetts.

Cell Signaling Technology: The life science company that life scientists built

Former Harvard Medical School neuroscientist Michael Comb founded Cell Signaling Technology (CST) because he saw a need for higher-quality antibody reagents in basic and clinical research. He built CST on a principle he inherited from his father: Research scientists are the best people to provide other researchers with the highest-quality products for their experiments.

"Working here has always felt to me like the science comes first, as opposed to being beholden to shareholders," says Katie Crosby, director of immunohistochemistry and the Western Blot Core, located at CST headquarters in Danvers, Massachusetts. Crosby has worked at CST since its beginning, when the life science company spun out from its parental powerhouse, New England BioLabs (founded by Michael Combs's father, Donald), in 1999. She's seen CST grow and come into its own in the last 23 years.

Headed by Michael Comb, a team of active research scientists like Crosby leads the family-owned, privately held company, with an ultimate commitment to advancing science in the right direction.

"What drives this company is the very deep desire to make an impact on biomedical research," says Roberto Polakiewicz, chief scientific officer of CST. "As a scientifically driven company, we develop products based on decisions made by scientists who understand the needs of researchers, because they understand the science."

CST also understands its scientist customers and the challenges of running rigorous, reproducible life science experiments. CST provides life sciences tools, such as antibodies used for Western blotting, immunohistochemistry, immunocytochemistry, flow cytometry, immunoprecipitation, and proteomics applications as well as antibody- and epigenetics-based kits. Company products focus on applications in cancer research, neurodegeneration, immunology and immuno-oncology, and epigenetics. Because its employees understand the high stakes involved in biomedical research, CST goes the extra mile to validate their antibody products. Their scientists even answer customers' support calls and emails and troubleshoot experiments to help researchers do their best work.

CST specializes in recombinant monoclonal antibodies, which offer several advantages over traditional monoclonal or polyclonal antibodies. Developing a recombinant antibody starts with knowing the specific genetic sequences for both the antibody's light chain and heavy chain, which can then be introduced into cultured cells for expression. Perhaps their biggest advantage is that recombinant antibodies offer excellent lot-to-lot consistency and can be produced in an almost unlimited supply, without the use of animals.

Validation campaigns as mini research projects

CST antibodies go through a rigorous validation process to ensure the antibody binds its target with high specificity in a variety of experimental scenarios.

"These validation campaigns are almost mini research projects," says Crosby. "We put in a tremendous amount of effort to show that the antibody identifies the right protein in the right application. That saves our customers from spending time doing that on their own."

PHOTO: COURTESY OF CST

To test their antibodies under various conditions, CST uses six complementary strategies that they call the "Hallmarks of Antibody Validation." These include testing antibodies across different model systems and cell types, with and without the target present, across a range of target concentrations, and in comparison to other antibodies that bind different sites, or epitopes, on the target protein. In this way, each antibody product can be recommended for high-specificity use under different experimental parameters. At the center of their validation approach is the conviction that no single assay can determine the validity of an antibody in a given application—and that includes CRISPR knockouts.

"We are particularly rigorous in our validation of antibodies," says Crosby. "When we put something out there for the community to use, we are confident in its performance."

Sometimes this in-depth validation work yields scientific discoveries all on its own. In an award-winning 2020 study of their monoclonal antibody products for two transcription coactivator proteins, yes-associated protein (YAP) and transcriptional coactivator with PDZ-binding motif (TAZ), CST scientists showed that these two closely related proteins could now be characterized separately in tissues using immunohistochemistry techniques (*1*). Before this, the two proteins implicated in some cancers were so similar in their sequences, domains, and expression that they were referred to as "YAP/TAZ," and scientists had difficulty distinguishing their functions.

"In some cancers we saw YAP and not TAZ, in others, the opposite. And in yet others, we saw them both together," says Crosby. "This suggested that they have individual roles in different cancers."

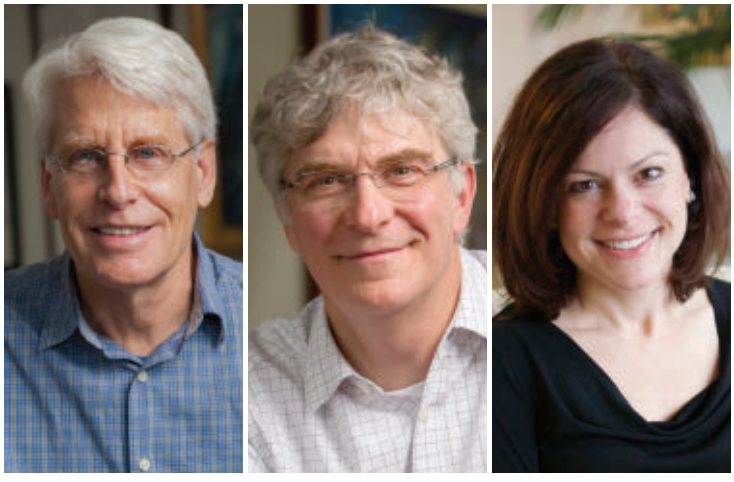
She notes that the stakes get much higher if researchers are trying to identify and characterize potential cancer biomarkers. "How do researchers interpret their data if they don't have confidence that the antibody is behaving in a very specific way?" asks Crosby.

CST is committed to using all available resources to push the science forward. Says CSO Polakiewicz: "Occasionally, a new product or technology can emerge from the research, but that is not the primary intent. Because CST scientists are scientists first, investigating intellectually interesting and impactful disease mechanisms helps keep them engaged with the scientific community and at the forefront of the field." CST scientists have published nearly 160 peer-reviewed papers since the company's founding (11 in 2021 alone), including contributing to research on COVID-19 through a recent collaboration with Weill Cornell Medicine about the impact of the drug alectinib on SARS-CoV-2 infection (*2*).

Science-based customer support

Because the scientists at CST spend time doing in-depth experiments to validate their products, they are the best resource when customers need technical support. The CST scientists who developed and tested the products answer customer phone calls and emails, will send supporting data to customers, and have even been known to run experiments to crack the thorniest customer queries.

"Technical Support at CST is definitely unique to the industry," says Craig Thompson, senior vice president of global operations. CST scientists not only understand the products and their applications, but often they are also familiar with the customers' biological systems. "The customer gets highly qualified support from a scientist who can troubleshoot the experiment technically and often provide guidance on experimental design."



Left: Michael Comb, **Center:** Roberto Polakiewicz, **Right:** Katie Crosby

Crosby notes that it's extremely satisfying to provide customers with this kind of detailed support to set them and their experiments on the right path to discovery.

Beyond the contributions to biomedicine, Crosby likes working at a company that also supports the broader community through sustainability initiatives and investments in science education and the arts. "We feel like our work is purposeful and we're being good citizens," she says.

Polakiewicz says that the company's scientific rigor, work ethic, and high-quality products and customer service stem from its family-owned structure and its deeply ingrained corporate social responsibility values. "Privately held companies like CST have the benefit of the long-term view without the pressures of the market," he says. That allows CST employees the space to do their own research and innovation and to lead the company in a researcher-centric way.

"From that culture comes the investment in doing good science, when we validate tools that other scientists will buy and use," says Polakiewicz. "That leads to an uncompromising business ethic that enables more good science to be done. We've been that way since we started, and that's never going to change."

References

1. K. Crosby *et al.*, *J. Histotechnology* **43**, 182–195 (2020).
2. T. M. Yaron *et al.*, *bioRxiv*, https://www.biorxiv.org/content/10.1101/2020.08.14.251207v4 (2020).

Sponsored by

Cell Signaling TECHNOLOGY®

PHOTO: COURTESY OF CST

HUMAN VACCINES PROJECT
DECODING THE IMMUNE SYSTEM

MICHELSON
MEDICAL RESEARCH
FOUNDATION

APPLY NOW
MICHELSON PRIZES:
NEXT GENERATION GRANTS

The Michelson Prizes: Next Generation Grants support young investigators applying disruptive research concepts to significantly advance the development of vaccines and immunotherapies for major global diseases.

A rigorous and competitive international search will identify the most innovative projects. We encourage applications from the full spectrum of scientific disciplines related to human immunology, vaccine, and immunotherapy research.

Are you 35 or younger?
Apply now for the 2022 Michelson Prizes!

humanvaccinesproject.org/michelson-prizes
#michelsonprizes

Dr. Rong Ma
Michelson Prize Winner 2021
Emory University

"As an early career researcher, the support from the Michelson Prize will give me a lot of independence and freedom to focus on my research."

APPLY TODAY

APPLICATIONS DUE:
JUNE 26, 2022

GRANT AWARD:

$150,000

CONTENTS

15 APRIL 2022
VOLUME 376
ISSUE 6590

234

Researchers in Thailand work late into the night to sample *Rhinolophus* bats that the team earlier found were infected with a virus related to SARS-CoV-2.

PHOTO: LAUREN DECICCA

NEWS

IN BRIEF

224 News at a glance

IN DEPTH

227 Global project gears up to study vaccine safety
Pandemic propels international efforts to understand incidence of rare side effects
By J. Couzin-Frankel

228 Earth's oldest land ecosystem spotted in drilled cores
Campaign probes for earliest signs of oxygen-producing life
By P. Voosen

230 Thermal batteries could back up green power
Efficiency jump in key component raises hopes for storing renewable energy as heat
By R. F. Service

231 How a site peddles author slots in reputable publishers' journals
Advertisements on Russian website promised to add names to articles that appeared in dozens of journals *By D. Singh Chawla*

232 Bills to ease cannabis research advance in U.S. Congress
Legislation may allow study of edibles or let universities grow their own plants for research *By M. Wadman*

233 Tiny labmade motors are poised to do useful work
Molecular-scale pumps that mimic the body's miniature machines could clean air and harvest precious metals *By R. F. Service*

FEATURES

234 Looking for trouble
Trapping bats with Supaporn Wacharapluesadee, who hunts for viruses to understand and prevent pandemic threats *By J. Cohen*

INSIGHTS

PERSPECTIVES

240 The fascinating world of biogenic crystals
The diverse properties of these crystals may lead to a variety of applications *By J. Prywer*
REPORT p. 312

242 Fullerenes make copper catalysis better
Ethylene glycol can be reliably produced by mild hydrogenation of dimethyl oxalate
By E. Gravel and E. Doris
REPORT p. 288

243 Regulatory CD8⁺ T cells suppress disease
A subset of CD8⁺ T cells regulate chronic inflammation by killing pathogenic CD4⁺ T cells *By A. Levescot and N. Cerf-Bensussan*
REPORT p. 265

245 Strong and fast hydrogel actuators
Plant cells inspire a hydrogel actuator that achieves ultrastrong and fast actuation
By Z. Jiang and P. Song
REPORT p. 301

246 Citizen science for studying earthquakes
Seismologist-citizen partnership helped understand the 2021 Haiti earthquake
By C. von Hillebrandt-Andrade and E. Vanacore
REPORT p. 283

247 Complex regulation of fatty liver disease
Hepatic lipogenesis is fine-tuned by mechanistic target of rapamycin (mTOR) signaling *By H. N. Ginsberg and A. Mani*
REPORT p. 264

248 How the gut talks to the brain
Peptidoglycans from gut microbiota modulate appetite through hypothalamic circuits *By A. Adamantidis*
REPORT p. 263

POLICY FORUM

250 Getting genetic ancestry right for science and society
We must embrace a multidimensional, continuous view of ancestry and move away from continental ancestry categories
By A. C. F. Lewis et al.

BOOKS *ET AL.*

253 Confronting climate injustice
Social, racial, and economic disparities are crucial considerations in climate policies
By M. Aczel

NOMIS & Science
YOUNG EXPLORER AWARD 2022

Research at the intersection of the social and life sciences

Unconventional. Interdisciplinary. Bold.

The **NOMIS & *Science* Young Explorer Award** recognizes and rewards early-career M.D., Ph.D., or M.D./Ph.D. scientists that perform research at the intersection of the social and life sciences. Essays written by these bold researchers on their recent work are judged for clarity, scientific quality, creativity, and demonstration of cross-disciplinary approaches to address fundamental questions.

A cash prize of up to 15 000 USD will be awarded to essay winners, and their engaging essays will be published in ***Science***. Winners will also be invited to share their work and forward-looking perspective with leading scientists in their respective fields at an award ceremony as well as a meeting of the NOMIS Board of Directors to consider future funding.

Apply by May 15, 2022

at www.science.org/nomis

NOMIS FOUNDATION
Creating the Spark

Science
AAAS

254 The complexity of chronic pain
A physician confronts an elusive physical phenomenon *By M. L. Meldrum*

LETTERS

255 Retraction
By K. Wagner et al.

255 Save Sri Lankan wildlife from foreign smugglers
By T. S. Priyadarshana

256 Scientists' right to speak to the press
By K. Foxhall et al.

256 Let's not abandon Russian scientists
By J. Holdren et al.

257 Technical Comment abstracts

RESEARCH

IN BRIEF

258 From *Science* and other journals

REVIEW

261 Ecosystem ecology
Evidence, causes, and consequences of declining nitrogen availability in terrestrial ecosystems *R. E. Mason* et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABH3767

RESEARCH ARTICLES

262 Lipidomics
Sphingolipids control dermal fibroblast heterogeneity *L. Capolupo* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABH1623

263 Neuroimmunology
Bacterial sensing via neuronal Nod2 regulates appetite and body temperature *I. Gabanyi* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABJ3986
PERSPECTIVE p. 248

264 Signal transduction
Inhibition of nonalcoholic fatty liver disease in mice by selective inhibition of mTORC1 *B. S. Gosis* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABF8271
PERSPECTIVE p. 247

265 Coronavirus
KIR^+CD8^+ T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19 *J. Li* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABI9591
PERSPECTIVE p. 243

266 Neuroscience
Compartmentalized dendritic plasticity during associative learning *S. d'Aquin* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABF7052

267 Neuroscience
Dynamic compartmental computations in tuft dendrites of layer 5 neurons during motor behavior *Y. Otor* et al.

REPORTS

276 Organic chemistry
Allylic C–H amination cross-coupling furnishes tertiary amines by electrophilic metal catalysis *S. Z. Ali* et al.

283 Natural hazards
Citizen seismology helps decipher the 2021 Haiti earthquake *E. Calais* et al.
PERSPECTIVE p. 246

288 Catalysis
Ambient-pressure synthesis of ethylene glycol catalyzed by C_{60}-buffered Cu/SiO_2 *J. Zheng* et al.
PERSPECTIVE p. 242

293 Aerosol optics
Amplification of light within aerosol particles accelerates in-particle photochemistry *P. Corral Arroyo* et al.

297 Cell biology
Epithelial monitoring through ligand-receptor segregation ensures malignant cell elimination *G. de Vreede* et al.

301 Hydrogels
Hydrogel-based strong and fast actuators by electroosmotic turgor pressure *H. Na* et al.
PERSPECTIVE p. 345

308 3d printing
Volumetric additive manufacturing of silica glass with microscale computed axial lithography *J. Toombs* et al.

312 Materials science
Complex morphologies of biogenic crystals emerge from anisotropic growth of symmetry-related facets *E. M. Avrahami* et al.
PERSPECTIVE p. 240

DEPARTMENTS

223 Editorial
Will ARPA-H work? *By H. H. Thorp*

318 Working Life
Listen to your body *By M. Zheng*

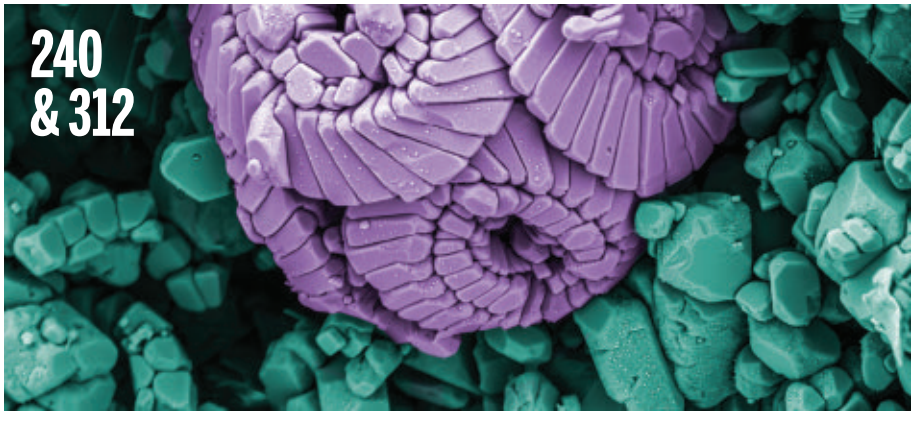
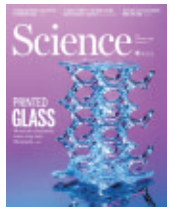


IMAGE: NANO CREATIVE/SCIENCE SOURCE

ON THE COVER

A glass structure about 4.5 mm tall with features as small as 0.25 mm is 3D printed with microscale computed axial lithography followed by high-temperature sintering. The process enables the synthesis of highly transparent and inert glass parts with fine details, which are useful for a variety of applications. See page 308. *Photo: Adam Lau/ Berkeley Engineering*



Science Careers 317

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2022 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2212; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.science.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Rutgers University: R&D with Impact

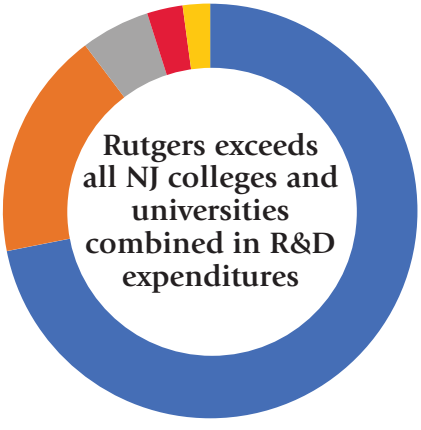
Committed to research and development driven by a relentless pursuit of excellence

$907.9 million
Research grants and sponsored programs

$720 million
Research and development expenditures

1187
Active technologies to license and market

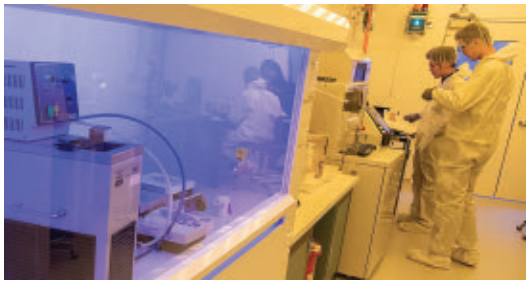
1000+
Companies partner with Rutgers researchers



Top 5 Direct Federal Sponsors of Rutgers Research

- National Institutes of Health $229.7M
- National Science Foundation $56.9M
- U.S. Health Resources and Services Administration $17.2M
- U.S. Department of Education $9M
- National Oceanic and Atmospheric Administration $6.8M

Top 100
World's Most Innovative Universities —Reuters



Recent Outcomes with Global Impact

- Invented COVID-19 diagnostic tests and serving as COVID-19 vaccines clinical trials site
- Pioneering use of underwater gliders for hurricane intensity forecasting
- Lead authorship of Intergovernmental Panel on Climate Change Reports
- Invented clean manufacturing technology for polymers and other monomers
- Co-lead investigator of large-scale urban test site for next-gen Wi-Fi

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

This is what **Excellence** *looks like*

excellence.rutgers.edu

Annual data current as of January 2022

EDITORIAL

Will ARPA-H work?

A new federal agency—approved last month by the United States Congress—is already off to a rocky start. The Advanced Research Projects Agency for Health (ARPA-H), proposed by President Biden in 2021, aims to tackle the most intractable biomedical problems by funding innovative, high-risk, high-reward research and swiftly turning discoveries into treatments and cures. But Congress gave the agency a much smaller budget than sought by the administration—$1 billion over 3 years, a fraction of the $6.5 billion requested. And as happens whenever there is new money and a new federal agency, a political scrum has erupted over who should control ARPA-H. It is now expected to answer to both the National Institutes of Health (NIH) and the Department of Health and Human Services (HHS). If it is to deliver on its mission, ARPA-H needs to be an autonomous entity that approaches biomedical research in a way never done before by the federal government. The stakes are high: If ARPA-H fails to produce new clinical advances relatively quickly, it will erode trust in US science. It's time for clear thinking and action about what it will take to make ARPA-H successful.

ARPA-H is modeled on the Defense Advanced Projects Agency (DARPA), an independent agency that reports to the US Department of Defense and whose goal is to fund breakthrough technologies for national security. For more than 60 years, DARPA has successfully accelerated innovation to application. The question is whether ARPA-H can follow this model because biomedical research is so different from the work supported by DARPA. In biomedical research, lots of things work in vitro, some things work in cells, fewer work in mice, and only a precious few actually work in humans. There is also a huge gap between basic research and immediate application, known as the "valley of death," where innovations get stuck. It will be a major feat for ARPA-H to bridge this chasm, particularly because the ecosystem in which it has been administratively placed—the NIH—is not one with a track record for driving this kind of research directly.

The NIH has been wildly successful at funding basic research, and although this has led to important clinical advances, these have almost always been carried forward by the private sector. The National Center for Advancing Translational Science was launched at NIH a decade ago to improve the bench-to-bedside process. It has had some success in fostering application of repurposed drugs, but no obvious de novo medicines or technologies have come from the division. Nevertheless, Francis Collins, the former NIH director and current White House science adviser, and other officials, including current NIH interim director Lawrence Tabak, have argued that ARPA-H should be embedded within NIH and controlled by the NIH director. Their rationale is twofold: The aims of ARPA-H fall squarely within the NIH's mission to improve human health, and ARPA-H needs access to NIH biomedical expertise. Meanwhile, critics have said that ARPA-H's mission begins where NIH's ends. "Everyone I know from NIH gets very excited about science, which is an input, but not the output, not the objective, of an ARPA," said former DARPA director Arati Prabhakar. Prabhakar is correct—basic science is exciting in a way that is absent from the sometimes tedious and scientifically conservative process of shepherding drugs from the lab to approval.

> "If ARPA-H fails to produce new clinical advances relatively quickly, it will erode trust in US science."

Xavier Becerra, the secretary of HHS, decided to place ARPA-H under the auspices of the NIH, but the director would report to Becerra, not the NIH director. So the ARPA-H director will have to answer to one boss but get administrative support from another. And it is unlikely that Becerra and NIH leadership will agree: Collins and Tabak want to fully absorb ARPA-H, but Becerra told Congress that NIH's role was merely to be the back office, providing human resources, payroll, and information technology. Indeed, Becerra has said that ARPA-H will be in a different physical location from NIH, which should help establish the new agency's independence.

President Biden has touted ARPA-H as a way to tackle Alzheimer's disease, cancer, and diabetes. It's hard to name three more difficult problems in biomedicine. It probably would have been better to take all of the new money and put it into basic research, which has led to new translational advances. But that opportunity has passed and it's time to stop arguing over who controls ARPA-H and do what it takes to make it successful.

–H. Holden Thorp

H. Holden Thorp
Editor-in-Chief, *Science* journals.
hthorp@aaas.org;
@hholdenthorp

PHOTO: CAMERON DAVIDSON

10.1126/science.abq4814

NEWS

> “Health equity work to me is not separate work. It’s one of the major challenges in modern medicine.”
>
> **Kirsten Bibbins-Domingo,** a Black physician-researcher, commenting in *STAT* after being named editor-in-chief of *JAMA*, which became embroiled last year in a debate over racism in medicine.

IN BRIEF

Edited by **Jeffrey Brainard**

The fossilized embryo of a pterosaur was found in a deposit from the end of the dinosaurs’ reign.

PALEONTOLOGY

Fossils of dinos killed by asteroid unveiled

A site in North Dakota has yielded what researchers contend are the first-known fossils of dinosaurs whose deaths can be directly linked to an asteroid impact that caused a major extinction some 66 million years ago. Reported last week, the discoveries, which have not been peer reviewed, also include pieces of amber that the team claims preserve shards of the asteroid itself, flung from the impact site in Mexico’s Yucatán Peninsula. Speaking at NASA’s Goddard Space Flight Center, paleontologist Robert DePalma, now a graduate student at the University of Manchester, said his team recovered the amber from sedimentary layers thought to date to minutes or hours after the impact, found at a fossil-rich site known as Tanis. Inside the amber, researchers identified the mineralogical signature of a type of asteroid known as a carbonaceous chondrite—not a comet, as others have suggested. DePalma also presented two fossils from those layers, including a pterosaur embryo. The BBC filmed DePalma and colleagues’ work at Tanis for a documentary set to air this week. Some researchers support DePalma’s claims, whereas others remain skeptical until they see the evidence for themselves, *The New York Times* reported.

Fourth shot helps, for a while

COVID-19 | A fourth dose of Pfizer’s COVID-19 vaccine offers older people some protection against serious illness caused by the Omicron variant of SARS-CoV-2 but only a brief defense against being infected at all, researchers report. In a study published on 5 April in *The New England Journal of Medicine*, they describe data from Israel, where people 60 and older have been offered a second booster shot since early January. The team found that 4 weeks after their fourth shot of the messenger RNA vaccine, recipients were half as likely to be infected as people who had received only three shots. The second booster’s protection against infection, however, had waned almost completely after 8 weeks. But recipients were also less than one-third as likely to suffer from severe COVID-19 symptoms 4 weeks after the fourth shot, a defense that was still strong after 6 weeks, the longest period for which these data were available. The findings come as the BA.2 variant of Omicron increases COVID-19 cases in the United States and other countries. European health agencies said on 6 April they had observed no “substantial waning” of protection from severe COVID-19 in singly boosted people between ages 60 and 80 who are not immunocompromised, and so there was no “imminent need” for them to get a second booster dose.

Professor’s conviction questioned

LEGAL AFFAIRS | Franklin Tao, a chemical engineer at the University of Kansas, Lawrence, last week was found guilty of lying about his ties to a Chinese research institution—and then received reason to hope his conviction might be overturned. Instead of setting a date to sentence Tao, U.S. District Judge Julie Robinson took the unusual step of asking federal prosecutors to explain why they believe Tao intended to defraud two U.S. agencies that had funded his research. In June 2019, Tao was the first academic scientist arrested under the U.S. government’s China Initiative, although he was never accused of economic espionage, the

ILLUSTRATION: MAURICIO ANTON/SCIENCE SOURCE

Police remove a protester from Spain's parliament as scientists demanded quick action on climate change.

PHOTO: ALDARA ZARRAOA/GETTY IMAGES

POLICY

Scientists arrested in widespread climate protests

Researchers in Washington, D.C., and Los Angeles were arrested last week after chaining themselves to a fence around the White House and the front doors of a J.P. Morgan Chase building, respectively, as part of global actions protesting governments' failure to stop climate change. The demonstrations were organized by the coalition Scientist Rebellion and included a total of about 1000 participants in 25 countries, the group says. Demonstrators clashed with police in many locations, including in Madrid, where protesters in lab coats threw fake blood on the steps of the Spanish parliament. The protests were in response to the release this month of the Intergovernmental Panel on Climate Change's most recent report, which implored world leaders to quickly switch to carbon-free energy. "We're not joking, we're not lying, we're not exaggerating. This is so bad that we're willing to take this risk," NASA climate scientist Peter Kalmus tweeted before being arrested in Los Angeles.

purported target of the initiative. In fall 2021, another academic scientist, Anming Hu, was acquitted of similar charges of failing to disclose ties to China after the judge in that case rejected the government's claim that Hu had sought to cheat NASA. "There is a lot of commonality between that case and this one, factually," Robinson told lawyers before the jury in Tao's case reached its verdict. Tao has been on unpaid leave from the university since his arrest and faces mounting legal bills.

Elusive woodpecker seen again?

ORNITHOLOGY | A team of scientists last week presented new evidence that the ivory-billed woodpecker (*Campephilus principalis*), long feared extinct, persists in swampy forests of Louisiana. Other researchers have voiced skepticism; previous claims from other scientists have not been verified. Project Principalis has been searching for the species for 10 years, recently using drones and automated, ground-level cameras. In a preprint posted to the bioRxiv server, Steven Latta of the National Aviary and colleagues describe several photographs and videos of what they identify as multiple ivory-billed woodpeckers. "We are fully confident," he says. In September 2021, the U.S. Fish and Wildlife Service proposed declaring the giant woodpecker extinct—the last conclusive evidence is from 1944—and removing it from the Endangered Species List. But in January, the service reopened a comment period to gather further evidence. Meanwhile, the researchers are continuing their search, including for traces of DNA.

Black people still wary of research

PUBLIC OPINION | Fifty years after the infamous syphilis experiment in Tuskegee, Alabama, was exposed, Black Americans harbor cautious, nuanced views of medical research, according to a Pew Research Center survey released last week. One-third say medical researchers do a good job all or most of the time, and 46%—almost half—say they do some of the time. But more than half of the 3546 Black respondents described misconduct by medical researchers as a very or moderately big problem that hasn't gotten better. Black Americans were more likely than all U.S. adults to say they have heard a lot (49%) or a little (26%) about the syphilis study, run by the U.S. Public Health Service from 1932 to 1972. Researchers deliberately withheld treatment from Black men with the sexually transmitted disease, leading to a worsening of symptoms and preventable deaths among participants. Today, Black people volunteer to participate in clinical research at disproportionately low rates, which scientists attribute to both the Tuskegee study's legacy and modern-day racism among health care professionals.

Donated cells shrink blood tumors

CLINICAL RESEARCH | Donated immune cells, mixed in a dish with a molecule that helps them home in on blood cancer cells, caused striking improvement in most of 22 people with lymphoma who received experimental infusions, researchers said this week. The new treatment, reported at the annual meeting of the American Association for Cancer Research, is

simpler and less costly than therapies such as genetically engineered chimeric antigen receptor (CAR)-T cells. In the new approach, researchers at MD Anderson Cancer Center used a "bispecific" antibody to connect a different type of immune cell, natural killer (NK) cells, with a surface protein on Hodgkin's lymphoma cells, then infused the complexes into patients. Of the 22 lymphoma patients, none experienced serious side effects, and tumors shrank in 17 of 19 who could be evaluated. Although some tumors resumed growing, seven of 13 patients who received the highest dose of cells were still in remission after 5 to 11 months.

One shot of HPV vaccine is enough

INFECTIOUS DISEASES | A single dose of vaccine against human papillomavirus (HPV) protects children and teens against later incidence of cervical cancer as well as two doses do, a World Health Organization (WHO) panel said this week—a finding that could allow health workers to stretch vaccine supplies and boost the number of people inoculated. In 2019, only 15% of girls worldwide had received two doses. Boys also receive the vaccine because HPV is linked to other kinds of cancers, but girls should receive priority, WHO's Strategic Advisory Group of Experts on Immunization said. Sexually transmitted HPV causes more than 95% of cervical cancer, the fourth most common type of cancer in women globally; 90% of these women live in low- and middle-income countries.

Gun seizures don't curb injuries

PUBLIC HEALTH | A closely watched 2016 California law that allows courts to temporarily take guns from people deemed a significant danger to themselves or others did not reduce the rate of firearm injuries in one metropolitan county, according to a first-of-its-kind study published last week in *JAMA Network Open*. Researchers at the University of California, Davis, and colleagues compared rates of firearm injuries from assaults and self-inflicted violence from 2016 to 2019 in San Diego county with those in a weighted combination of 27 other counties that issued far fewer such orders in that period. They found no significant differences in firearms injuries, and speculate that people who committed assaults after losing their firearms may have obtained new ones illicitly. A separate study by Stanford University researchers found that Californians living with someone who lawfully owns a firearm were more than twice as likely to be murdered than those living with nonowners. For at-home homicides, the risk was seven times as high, and 84% of victims were women, according to the study of more than 17 million Californians tracked for up to 12 years. The study was published last week in the *Annals of Internal Medicine*.

BY THE NUMBERS

1895.7

The 2021 level of atmospheric methane in parts per billion (ppb), a new record. The level of the powerful greenhouse gas rose 17 ppb last year, the largest absolute increase since modern records began in 1983. (U.S. National Oceanic and Atmospheric Administration)

15.8%

Share of the world's population who report a headache each day, on average, making it one of the most common health problems, a meta-analysis found. (*The Journal of Headache and Pain*)

46%

Portion of U.K. parents surveyed who call school work in physics "complicated," which may be discouraging their children from studying the subject. (Institute of Physics)

A bumble bee hovers near a rapeseed plant, a crop commonly sprayed with the insecticide sulfoxaflor.

PHOTO: PHIL SAVOIE/NLP/MINDEN PICTURES

AGRICULTURE

Europe will limit leading pesticide to spare pollinators

Just 7 years after authorizing outdoor use of the insecticide sulfoxaflor as a less toxic alternative to other products, the European Commission said last week it plans to ban it. The move will prohibit farmers in the European Union from treating crops with the widely used pesticide, manufactured by Corteva Agriscience. Regulators cited risks to the common bumble bee (*Bombus terrestris*) and other species of pollinators living next to farm fields. Sulfoxaflor threatens bees when they collect pollen and nectar containing the chemical. Its use will still be allowed in greenhouses, which regulators expect will protect wild bees from exposure. The U.S. Environmental Protection Agency (EPA) does not allow sulfoxaflor on crops, such as cotton, that attract bees. But both EPA and European countries have allowed emergency use of sulfoxaflor and other insecticides if crops are threatened with a devastating infestation, which some environmental advocates say is a loophole that will continue to harm pollinators.

SCIENCE.ORG/NEWS
Read more news from *Science* online.

Mass vaccination efforts for COVID-19, like this one in Toronto, gave new life to a vaccine safety project that could cover hundreds of millions of people worldwide.

PHOTO: ZOU ZHENG/XINHUA/GETTY IMAGES

COVID-19

Global project gears up to study vaccine safety

Pandemic propels international efforts to understand incidence of rare side effects

By **Jennifer Couzin-Frankel**

Scientists in more than 20 countries, on every continent save Antarctica, have started to gather data for the largest ever vaccine safety project. Members of the effort, called the Global Vaccine Data Network (GVDN), fruitlessly sought funding after conceiving the project more than 10 years ago. But the mass vaccinations during the COVID-19 pandemic breathed new life into the project. With the ability to draw on data from more than 250 million people, the network will investigate rare complications linked to COVID-19 vaccines in hopes of improving prediction, treatment, and potentially prevention of these side effects.

"You really need global data in order to understand" rare vaccine side effects, says Gregory Poland, a vaccinologist at the Mayo Clinic. Poland, who's not involved in GVDN, himself developed severe tinnitus about 90 minutes after his second vaccine dose, which he suspects is related to the shot. Studying potential vaccine complications "is a very neglected area," he says.

Bruce Carleton, a clinical pharmacologist at the University of British Columbia, Vancouver, and head of the GVDN genomics effort, draws an analogy to air travel safety. There, improvements often came after ultrarare crashes. Airplanes, Carleton says, were shored up by "learning from those events, not denying them." With billions of doses of COVID-19 vaccines administered, it's clear the vaccines are "very safe for most people," he continues. At the same time, "There probably are patients that may, in fact, suffer harm."

Doing this research comes with steep scientific hurdles, among them the rarity of serious problems. The largest vaccine studies have included about 1 million people, and even that can be too small to nail down side effects. "If you had something that happened normally to one in 100,000 people, and you wanted to see if the vaccine doubled the risk, you'd need a study with about 4 million people," says Helen Petousis-Harris, a vaccinologist at the University of Auckland who jointly heads GVDN with Steven Black, a pediatric infectious disease specialist formerly at Cincinnati Children's Hospital.

The idea for GVDN came to Black around 2009, when the H1N1 flu pandemic hit and a mass vaccination campaign began. Some countries detected an increased risk of narcolepsy from the vaccine, called Pandemrix, but others did not. The variability might have reflected differences in vaccine surveillance, which varies geographically, with some countries relying on passive reporting and others combing through health records for patterns. Or, scientists came to suspect, the immune response to vaccination might somehow have interacted with flu infection to trigger the narcolepsy. Black thought consistent data could help solve such mysteries.

He set out to globalize vaccine safety research. Funding, however, was nowhere to be found. Then in 2019, the Bill & Melinda Gates Foundation offered seed money for a meeting. About 60 vaccine safety specialists descended on a lakeside village in France and GVDN was born. The network got a jumpstart with support from Petousis-Harris's university and Auckland UniServices Ltd., a nonprofit owned by the institution.

Carleton and Daniel Salmon, a vaccine safety researcher at Johns Hopkins University, conjured up its first project: a study of the potential risk of Guillain-Barré syndrome, a rare neurologic condition, from flu vaccines. In early February 2020, just as

COVID-19 was taking hold, they submitted a funding proposal to the National Institutes of Health (NIH), briefly noting that if a pandemic emerged and vaccines were developed, their network would be well-suited to studying any side effects. The application was turned down, as were subsequent applications to the World Health Organization and NIH.

Then in April 2021, with COVID-19 vaccination drives underway, the Centers for Disease Control and Prevention (CDC) awarded GVDN $5.5 million over 3 years to study the safety of the vaccines. It was a shoestring budget but enough to design several projects, each drawing on large health systems, regions within a country, or, in some cases, as in New Zealand, a nation's entire population. One will study heart inflammation associated with the messenger RNA vaccines from Pfizer and Moderna. Another will probe vaccine-induced thrombotic thrombocytopenia, a dangerous clotting disorder linked to viral vector vaccines made by AstraZeneca and Johnson & Johnson. The network will also examine the risk of Guillain-Barré syndrome after COVID-19 vaccination, among other projects.

"We've been trying for so long" to get projects like these started, says Robert Chen, scientific director of the Brighton Collaboration, which studies vaccine safety, and a former director of CDC's vaccine safety program. Chen notes that those who advocate for improved vaccine safety "have unfortunately been lumped into the antivaccine groups" at times, even though their goal, he says, is to make rare side effects even rarer.

Pulling the venture together has been "gnarly," Petousis-Harris says. It's meant middle-of-the-night conference calls across a dozen time zones, and months spent harmonizing the definition of a health condition, like myocarditis, across hospital systems and countries. Getting at the risk increase of complications after vaccination is also statistically complex. Some studies do this by comparing a vaccinated population with an unvaccinated one and assessing whether more of the former develop, say, myocarditis. But unvaccinated people differ from vaccinated ones in other ways, which may cloud the results. Instead, GVDN will use a method called a "self-controlled case series." For example, they will identify everyone who suffered myocarditis in the 60 days that followed their last vaccine dose. Statisticians can then examine whether and to what degree participants were more likely to contract the heart condition immediately after vaccination versus weeks later.

A massive data set could also begin to crack other mysteries—in particular, who's at risk. Is postvaccine myocarditis more likely, for example, if someone has another health condition or takes a certain medication?

Another dream is to understand the biology underpinning side effects. Insights could come from another nascent global effort, called the International Network of Special Immunization Services (INSIS), which is now finalizing funding agreements. Whereas GVDN aims to scoop up and analyze COVID-19 vaccine data worldwide and tackle genomics questions, INSIS will examine the biology and immunology of postvaccine problems as they are happening. The network's leader, Karina Top, a pediatric infectious disease specialist at Dalhousie University, is working with GVDN to identify patients and share data.

For now, the new efforts don't include difficult-to-diagnose health problems that may be linked to the vaccines. Some people have described Long Covid–like symptoms, such as chronic headaches and irregular heart rate and blood pressure, soon after vaccination, but studying this phenomenon is far more difficult (*Science*, 28 January, p. 364). Headaches, for instance, are so common that no vaccine surveillance system would detect an imbalance, says Rebecca Chandler, who works on vaccine safety at the Coalition for Epidemic Preparedness Innovations. She says that to discern patterns, it's crucial to review not just single words in vaccine safety reports, but narratives from doctors and patients.

"The subjective reports should not be discounted as meaningless or unrelated," Carleton says. GVDN and INSIS, he and others hope, will expand their scope as time goes on, funding permitting.

Carleton's suspicion is that much of a person's risk comes down to genetics. Through GVDN, he'll test whether certain gene variants raise risk of postvaccine complications. He's also planning a solo project, setting up a website for anyone who believes they've suffered a post–COVID-19 vaccine adverse event, inviting them to ship him a saliva sample and health records.

The GVDN team hopes to have initial data by late summer. "My sincere hope is that once we prove the value of something like the GVDN, we can secure longer term sustainable funding," says Ann Marie Navar, a cardiologist at the University of Texas Southwestern Medical Center and a researcher with GVDN. Chen, who's advising INSIS, is cautiously optimistic, but knows the road ahead could be bumpy. "Unless we find a way to stabilize" funding for these projects, he says, "it's very easy for them to just collapse." ■

EARTH SCIENCE

Earth's oldest land ecosystem spotted in drilled cores

Campaign probes for earliest signs of oxygen-producing life

By **Paul Voosen**

Beneath the Barberton Makhonjwa Mountains, home to South Africa's original gold rush, lies something more scientifically valuable than any precious metal: Earth's first land ecosystem, trapped in a 3.2-billion-year-old rock formation called the Moodies Group. In roadcuts and mineshafts, scientists had already glimpsed fossilized remnants of the slimy microbial mats thought to have covered the ancient rivers, beaches, and estuaries. Now, they are drilling into the terrain for the first time, retrieving fresh samples of what may have been Earth's first microbial producers of oxygen.

"It's really lucky there are places as old as this," says Tanja Bosak, a geobiologist at the Massachusetts Institute of Technology who is unaffiliated with the project. Although older signs of life have been found in South Africa and Australia—and potentially Greenland—in what were once ocean deposits, no other spots record primordial life on land so convincingly, she says. "This covers a not-well-understood time in Earth's history."

When the Moodies Group formed, Earth would have been nearly unrecognizable. Its atmosphere, rich in methane and carbon dioxide but nearly devoid of oxygen, kept the planet warm while the Sun was young and faint. Land was scarce because plate tectonics, the process that assembles continents, was just getting going. Here and there, however, volcanic archipelagos like the Moodies Group pierced the waters. Beaches ringing the volcanoes would have been ideal spaces for life to evolve and spread, says Christoph Heubeck, a sedimentary geologist at the Friedrich Schiller University of Jena. He leads the $2 million Barberton Archaean Surface Environments (BASE) project, which plans to complete drilling its eighth and final core next month.

A drilling operation in South Africa's Barberton Makhonjwa Mountains extracted 3.2-billion-year-old cores that hold some of Earth's earliest terrestrial life.

PHOTO: CHRISTOPH HEUBECK

The cores the team has already extracted, from deposits 200 meters below the surface, are rich in fossilized slimes. "We've drilled through hundreds of meters of them," Heubeck says. Their nature, however, is a mystery.

Other ancient microbial fossils in the Moodies Group, found in what were marine and subsurface deposits, probably fed on sulfates or used a primitive form of photosynthesis to feed on iron. But those metabolic pathways would not have worked well in the Sun-soaked shallow waters in which the slimes lived. Heubeck believes these microbes were early ancestors of cyanobacteria, which some 800 million years later flooded the atmosphere with oxygen in what's called the Great Oxidation Event. "The production of oxygen appears to be a process invented early in Earth's history," he says.

It's a controversial claim. If oxygen-producing photosynthesis had evolved so early, some researchers argue, the Great Oxidation Event would have promptly followed. But evidence for early "oxygen oases" has grown. Geochemists have found mineral deposits from well before the Great Oxidation Event that needed oxygen to form. And genetic analysis of cyanobacteria suggests they evolved, on land, around the same time as the Moodies Group, says Patricia Sanchez-Baracaldo, a paleobiologist at the University of Bristol who is unaffiliated with BASE. "The genomic record is independent and consistent with the idea that those were early ancestors of cyanobacteria."

Heubeck and colleagues hope the fresh, unaltered microbial mats in the cores will yield decisive evidence: geochemical traces of oxygen production that have been missing in previous, exposed samples. That hunt will begin in earnest later this year, when the team begins to pore over half of the cores at a "sampling party" in Germany; the other half will remain in South Africa as an archive.

The cores could contain other scientific treasures. In 2010, Emmanuelle Javaux, an astrobiologist at the University of Liège, reported finding walled spherical microbial fossils up to 300 micrometers in diameter, hundreds of times the size of a typical bacterium, in mudstones extracted from a gold mine in the Moodies Group. Some thought the jumbo microbes were the world's oldest eukaryotes—organisms with complex cells like our own—by 1 billion years, but confirmation proved elusive. Javaux hopes the BASE cores will capture the same fossils in better condition. "Now we just have to find them," she says.

The BASE cores could also hold clues to the climate of that ancient landscape. One core contains what appears to be lithified layers of soil, which could capture indicators of the atmosphere's composition. Offshore shales may record how the islands' volcanic basalt eroded. Whether it broke off in chunks, as happens in today's Arctic, or was ground down into bits as in tropical climates could hint at the ancient temperatures. Other samples capture an interwoven pattern of sand and mud layers, assembled by the ancient tides. The Moon was much closer to Earth at the time, and the tidal record could pin down its distance.

The cores should also contain a record of lightning strikes, which create strong magnetic fields that can be imprinted on rocks. Lightning might have supplied a key nutrient to the ancient ecosystem by splitting apart the tough molecular bonds of atmospheric nitrogen, enabling the atoms to form the compounds that life depends on. Because the microbes that break down nitrogen today were scarce or even nonexistent, the strike rate alone would reveal how much of this important nutrient was being added to the surface. "This nitrogen flux is potentially a major component of the biosphere at the time," says Roger Fu, a planetary scientist at Harvard University.

In many ways, the Moodies Group cores are preparing geologists for the work to come when rock samples are returned from another 3-billion-year-old terrain—on the surface of Mars. Later this month, NASA's Perseverance rover will reach a fossilized river delta and begin to drill cores. If, as hoped, future Mars missions return those cores to Earth, the lab techniques used on the BASE cores will come in handy, Bosak says. "Looking at these well-preserved sediments on Earth will tell us what the ideal case will be from Mars." ■

ENERGY TECHNOLOGY

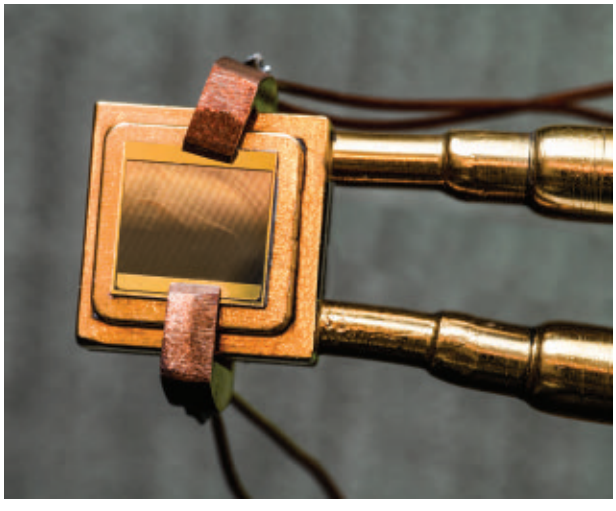
Thermal batteries could back up green power

Efficiency jump in key component raises hopes for storing renewable energy as heat

By **Robert F. Service**

How do you bottle renewable energy for when the Sun doesn't shine and the wind won't blow? That's one of the most vexing questions standing in the way of a greener electrical grid. Massive battery banks are one answer. But they're expensive and best at storing energy for a few hours, not for days long stretches of cloudy weather or calm. Another strategy is to use surplus energy to heat a large mass of material to ultrahigh temperatures, then tap the energy as needed. This week, researchers report a major improvement in a key part of that scheme: a device for turning the stored heat back into electricity.

A team at the Massachusetts Institute of Technology (MIT) and the National Renewable Energy Laboratory achieved a nearly 30% jump in the efficiency of a thermophotovoltaic (TPV), a semiconductor structure that converts photons emitted from a heat source to electricity, just as a solar cell transforms sunlight into power. "This is very exciting stuff," says Andrej Lenert, a materials engineer at the University of Michigan, Ann Arbor. "This is the first time [TPVs have] gotten into really promising efficiency ranges, which is ultimately what matters for a lot of applications." Together with related advances, he and others say, the new work gives a major boost to efforts to roll out thermal batteries on a large scale, as cheap backup for renewable power systems.



A thermophotovoltaic cell turns furnacelike heat into electricity.

The idea is to feed surplus wind or solar electricity to a heating element, which boosts the temperature of a liquid metal bath or a graphite block to several thousand degrees. The heat can be turned back into electricity by making steam that drives a turbine, but there are trade-offs. High temperatures raise the conversion efficiency, but turbine materials begin to break down at about 1500°C. TPVs offer an alternative: Funnel the stored heat to a metal film or filament, setting it aglow like the tungsten wire in an incandescent light bulb, then use TPVs to absorb the emitted light and turn it to electricity.

When the first TPVs were invented in the 1960s, they only converted a few percent of the heat energy into electricity. That efficiency jumped to about 30% in 1980, where it has largely been stuck ever since. One reason is that tungsten and other metals tend to radiate photons across a broad spectrum, from high-energy ultraviolet to low-energy far-infrared. But all photovoltaics—TPVs included—are optimized to absorb photons in a narrow range, meaning light with higher and lower frequencies tends to be wasted.

For the new device, Asegun Henry, an MIT mechanical engineer, tinkered with both the emitter and the TPV itself. Previous TPV setups heated the emitters to about 1400°C, which maximized their brightness in the wavelength range for which TPVs were optimized. Henry aimed to push the temperature 1000°C higher, where tungsten emits more photons at higher energies, which could improve the energy conversion. But that meant reworking the TPVs as well.

With researchers at the National Renewable Energy Laboratory, Henry's team laid down more than two dozen thin layers of different semiconductors to create two separate cells stacked one on top of another. The top cell absorbs mostly visible and ultraviolet photons, whereas the lower cell absorbs mostly infrared. A thin gold sheet under the bottom cell reflects low-energy photons the TPVs couldn't harvest. The tungsten reabsorbs that energy, preventing it from being lost. The result, the group reports this week in *Nature*, is a TPV tandem that converts 41.1% of the energy emitted from a 2400°C tungsten filament to electricity.

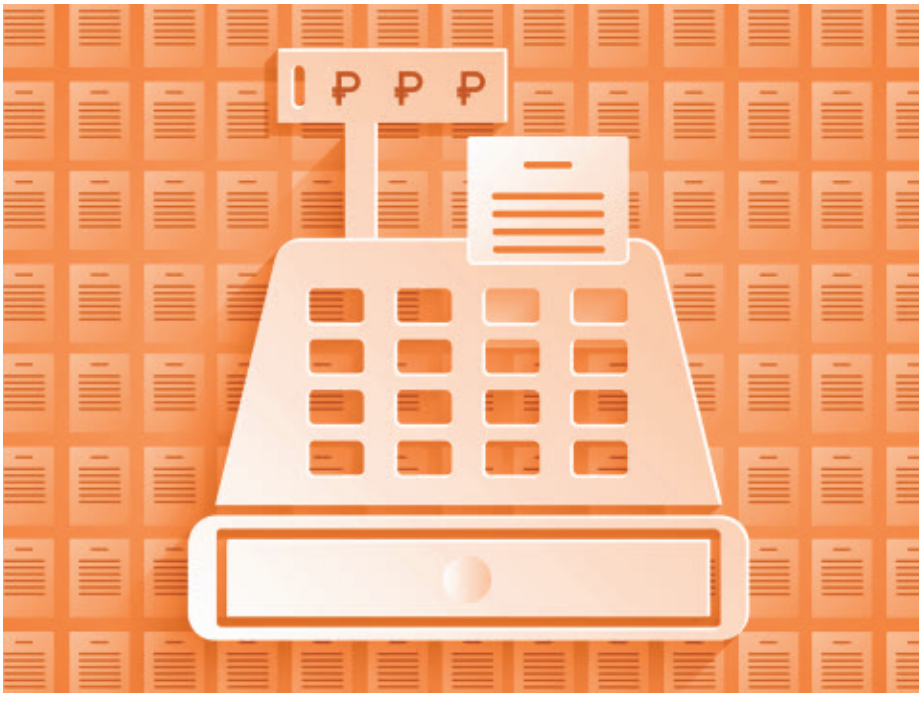
Henry's team sees ways to do even better. In the 8 October 2020 issue of *Nature*, Lenert and his colleagues reported a mirror able to reflect nearly 99% of unabsorbed infrared photons back into the heat source. Coupling the mirror with the MIT group's improved TPVs could yield another big boost. "We think we have a clear path to 50% efficiency," Henry says.

The TPVs are made from III-V semiconductors, named for where their component elements fall in the periodic table, which are more expensive than the silicon used in rooftop solar cells. But other parts of a thermal battery, including graphite, are cheap. In a 2019 paper, Henry and his colleagues had calculated that even a 35% efficiency in heat-to-electricity conversion would make the technology economically viable. The team has also created ceramic pumps that can handle the ultra–high-temperature liquid metals needed to carry heat around an industrial scale heat energy storage setup. "They've built a foundation for storing and converting heat at those high temperatures," Lenert says.

This progress has triggered commercial interest. Antora Energy in California launched a thermal energy company in 2016. Lenert and others are eyeing their own startups. And Henry recently launched a venture—Thermal Battery Corp.—to commercialize his group's technology, which he estimates could store electricity for $10 per kilowatt-hour of capacity, less than one-tenth the cost of grid-scale lithium-ion batteries. "Storing energy as heat can be very cheap," even for many days at a time, says Alina LaPotin, an MIT graduate student and first author of the current *Nature* paper.

Henry and others add that thermal storage systems are modular, unlike fossil fuel plants, which are most efficient at a massive, gigawatt scale. "That makes them equally good at providing power for a small village or a large power plant," says Alejandro Datas, an electrical engineer at the Polytechnic University of Madrid—and for storing power from solar and wind farms of any size. "This is the beauty." ■

PHOTO: FELICE FRANKEL



PUBLISHING

How a site peddles author slots in reputable publishers' journals

Advertisements on Russian website promised to add names to articles that appeared in dozens of journals

By **Dalmeet Singh Chawla**

Since 2019, Anna Abalkina has been monitoring a website that offers an illicit way for scientists to burnish their CVs. The site, operated from Russia, openly offers to sell authorship slots on soon-to-be published scientific papers, for fees ranging from several hundred dollars to nearly $5000.

Abalkina, a sociologist at the Free University of Berlin, has documented what appears to be a flourishing business on the site, www.123mi.ru. Since it debuted in December 2018, she has analyzed more than 1000 advertisements posted there and found at least 419 that seemed to match manuscripts that later appeared in dozens of different journals, she reported in a preprint posted on arXiv in March.

More than 100 of the papers she identified were published in 68 journals run by established publishers, including Elsevier, Oxford University Press, Springer Nature, Taylor & Francis, Wiley-Blackwell, and Wolters Kluwer, although most of these were specialized publications. Russian authors outnumbered any other nationality on the website's tally of recent contracts.

Run by International Publisher LLC, the site is one of many illicit "paper mills" that leaders in scientific publishing worry are increasingly corrupting the literature by selling bogus authorship or prewritten papers. But its scale and brazenness are unusual, as are the insights Abalkina has gleaned into its workings.

Her findings are "fascinating," says Elisabeth Bik, an independent science integrity expert in San Francisco who believes they reflect fallout from Russia's 2012 decision to set policies tying researchers' promotions and financial rewards to their volume of scholarly publications.

"It is another example of what can go wrong in scientific publishing if the pressure to publish is increased," says Bik, who has studied paper mills based in China.

To lure prospective customers, the advertisements on www.123mi.ru provide tantalizing details about each paper, which it claims are already accepted for publication. They include its topic, the number of authors, and sometimes its abstract. The advertisements also provide hints about the prestige and impact of the journal in which the paper will appear, including whether it is indexed in the Scopus and Web of Science databases.

Prices for authorship slots vary depending on their position in the authors list and the impact factor of the journal, Abalkina found. Costs have varied from about 15,000 rubles ($175) to 410,000 rubles ($4800), with first author slots usually the most expensive. Based on these fees, Abalkina estimates that from 2019 to 2021, International Publisher raked in about $6.5 million. (The website does not specify how much its customers actually paid.) To keep the deals hush-hush, the contract includes a confidentiality clause.

The advertisements withhold the name of the journal, which the purchaser is told only after paying the fee. Abalkina's paper quotes claims by the website that it has split its fees with some unidentified journals to ensure their participation in the scheme.

Several of the largest publishers of journals identified in Abalkina's study—Oxford University Press, Springer Nature, Taylor & Francis, and Wiley-Blackwell—say they are examining papers Abalkina brought to their attention. Elsevier has procedures to detect authorship changes after a manuscript is submitted, which editors must approve, a spokesperson said. As for the papers identified by Abalkina, "there were few indications of authorship changes occurring after submission."

Chris Graf, director of research integrity at Springer Nature, declined to discuss specifics but called paper mills "bad for both the research and the publishing communities. Alongside investigating individual cases and retracting compromised papers, we've been reviewing our processes and making investments in technologies to help us identify attempts to manipulate our systems."

Science contacted International Publisher—which says it is headquartered in Moscow and has offices in Ukraine, Kazakhstan, and Iran—for comment multiple times by email, phone, and WhatsApp but received no response. Its chief editor, according to her LinkedIn page, is Ukraine-based philologist Ksenia Badziun, who says she graduated from Taras Shevchenko National University of Kyiv. The company has continued to operate during Russia's invasion of Ukraine.

Science also contacted 20 corresponding authors of papers identified by Abalkina; most did not respond. One, who asked not to

IMAGE: VALERII MINHIROV/ISTOCK.COM; ALIAKSEI BROUKA/ISTOCK.COM; OKSANA SAZHNIEVA/ISTOCK.COM. ADAPTED BY C. AYCOCK/*SCIENCE*

be named, said he knows nothing about International Publisher or its activities and that all listed co-authors contributed to the work.

Kim-Hung Pho, a statistician at Ton Duc Thang University who co-authored two of the flagged studies—both published by *Digital Scholarship in the Humanities*, run by Oxford University Press—also told *Science* he has no knowledge of www.123mi.ru. "I don't have any funds to do scientific research, so I have absolutely no money to buy [authorship in] these articles, and there is no pressure to do this."

In 2021, publishers retracted a record 724 articles traced to paper mills, part of a grand total of more than 1000 such articles retracted during the past decade, according to a database maintained by the Retraction Watch website. (More than 4 million scholarly papers are now published annually.)

At least two nonprofit groups that advocate for honest practices in publishing have issued guidance to journal editors about how to deter purchased authorships. The Committee on Publication Ethics and the International Committee of Medical Journal Editors recommend that editors require authors who request to add an author after submitting a manuscript to provide an explanation and signed permission from all other listed authors.

But some observers suggest journal editors should do more. If they discover papers with authors who paid to be listed, they should flag them by attaching a written "expression of concern" because "any association with the mill raises some question about the integrity of the paper," says Bryan Victor, who studies social work at Wayne State University. In December 2021, he co-authored a separate analysis of www.123mi.ru on Retraction Watch, which described nearly 200 papers that may match authorships advertised there; subsequently he and a colleague posted a catalog of contracts displayed on the site, which together refer to about 1500 articles.

How editors could identify fraudulent authors before publication remains murky. But the papers do offer clues, Abalkina found. For example, a few list authors in multiple unrelated academic departments, making it unlikely they collaborated. In other cases, the authors' specialties don't match the manuscript's title.

To avoid scrutiny from editors, International Publisher appears to avoid repeatedly targeting the same journals, Abalkina says. "That makes it impossible for an editor to detect some anomalies," she says. ■

Dalmeet Singh Chawla is a science journalist in London.

SCIENCE POLICY

Bills to ease cannabis research advance in U.S. Congress

Legislation may allow study of edibles or let universities grow their own plants for research

By **Meredith Wadman**

For decades, marijuana's federal criminal status has raised obstacles for researchers trying to study the plant's health impacts, such as forcing them to rely on a single grower or keep samples in 340-kilogram safes under 24/7 surveillance. Now, the U.S. Congress is poised to eliminate some of those hurdles.

Two bills, passed unanimously in the Senate late last month and by an 82% vote in the House of Representatives last week, would ease storage rules, streamline application procedures for would-be cannabis researchers, and allow them to modify research protocols more easily. The Senate version allows universities to grow their own plants for research, and the House bill allows researchers to study products sold at dispensaries in states that have legalized marijuana.

Researchers need to study "these chemicals that the rest of the world can get very easily."

Ziva Cooper, University of California, Los Angeles

The two bills share a common theme, says Representative Andy Harris (R–MD), a leading co-sponsor: giving researchers "access to a wider range of products that reflect the modern marijuana marketplace." He predicts that House and Senate negotiators will iron out the differences "by summer or fall" and the resulting measure will become law. "My hope is that [in facilitating research] we once and for all determine whether marijuana or its components actually have a role in the treatment of disease," he says.

The United States designates marijuana as a schedule I drug, in the same category as heroin and LSD. Researchers studying it must win Drug Enforcement Administration (DEA) registration, an arduous process, and keep samples in high-security vaults or safes. Until recently, they could only study plants from one grower, the University of Mississippi, although DEA recently registered a few additional growers.

As a result, "All that researchers can get is cannabis that's worse than probably what my parents smoked in college," says Logan Leichtman, an attorney for cannabis companies in Portland. Oregon. Plants from the Mississippi farm have far less delta-9-tetrahydrocannabinol (THC), cannabis' main psychoactive ingredient, than many products legally sold today for medical use in 37 states and Washington, D.C.

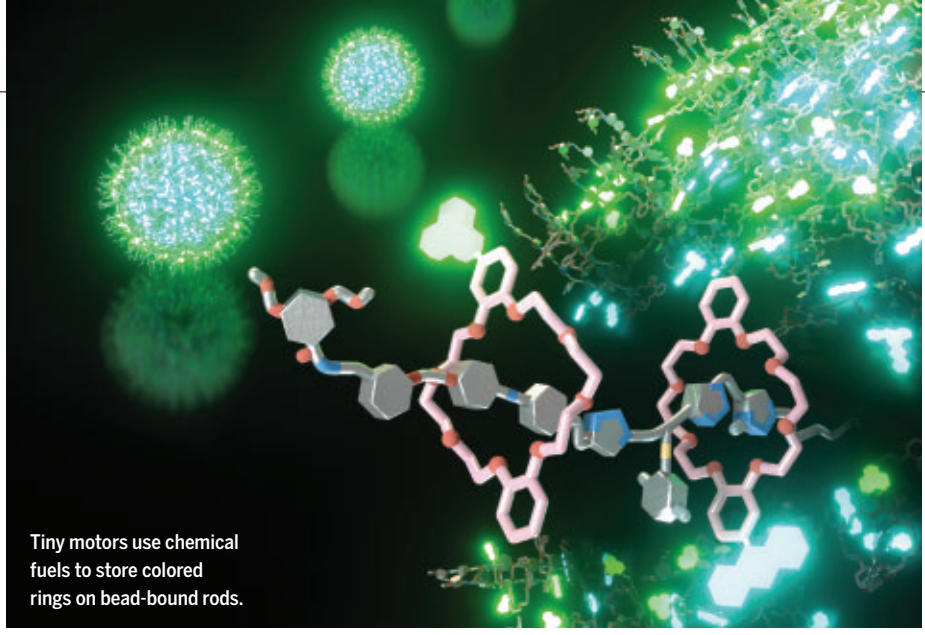
Despite the broad and growing medical usage, the U.S. Food and Drug Administration (FDA) has only approved a few cannabis-based therapies, including two synthetic THC-based medicines for nausea and loss of appetite and one drug, cannabidiol (another chemical in cannabis), for intractable epilepsies. Researchers are keen to examine cannabinoids as potential therapies for chronic pain, cancer, anxiety, and other conditions.

To do that, "Researchers need to be able to study these chemicals that the rest of the world can get very easily," says Ziva Cooper, director of the Cannabis Research Initiative at the University of California, Los Angeles.

The House provision allowing access to dispensary products may not survive. The more modest Senate version "has been very carefully negotiated to make it able to pass by unanimous consent through the Senate, which is no small feat," says a policy staffer for Senator Dianne Feinstein (D–CA), the bill's lead sponsor. Such unanimity could speed final passage. "Our hope is that the House will consider adopting a bill that looks a lot like ours."

The Senate bill, which has the backing of the American Medical Association, would also allow researchers to import cannabis and require DEA to register companies producing FDA-approved drugs.

Still, studying marijuana will remain challenging because of funding constraints and its continued U.S. criminal status, scientists say. "Symbolically I think [the bills] are really important," Cooper says. "But for somebody who does human research … [studying cannabis is] still going to be hard." ■



Tiny motors use chemical fuels to store colored rings on bead-bound rods.

IMAGE: ANNA TANCZOS/SCICOMM STUDIOS

CHEMISTRY

Tiny labmade motors are poised to do useful work

Molecular-scale pumps that mimic the body's miniature machines could clean air and harvest precious metals

By **Robert F. Service**

Tiny molecular machines make life possible. Spinning rotary motors generate the chemical fuel our tissues need, miniature walkers carry nutrients around cells, and minute construction crews build proteins. Now, chemists are building their own smaller and simpler versions of these biological machines, hoping to harness their powers.

The puny molecular pumps and motors scientists describe in three recent studies are not quite ready to make their real-world debut. But the work shows it's possible to get teams of motors all working in the same direction to concentrate target chemicals in a confined space—the same feat that protein motors perform in our cells. The feat raises hopes that future motors could suck carbon dioxide from the air and harvest valuable metals from seawater.

"These are very important steps toward useful real-life molecular machines," says Ivan Aprahamian, a chemist at Dartmouth College who wasn't involved with the recent set of studies.

It isn't easy for a molecular-scale motor to do useful work. Because of the vanishingly small size of molecules, a chemical reaction that causes a molecular rotor to spin clockwise is equally likely to spin it counterclockwise. And heat jostles molecules randomly in all directions. "At such small scales, random chaotic motion of components and molecules is inevitable," says Nathalie Katsonis, a chemist at the University of Groningen.

Fraser Stoddart has been working to overcome this challenge for years. The organic chemist at Northwestern University created some of the world's first small, chemical-based molecular machines, sharing a Nobel Prize for his research in 2016. His team designed rings that would slip on and off a molecular axle in response to different chemicals. But because those machines drifted around randomly in solution, collections of them didn't coordinate their tasks in any particular direction, which meant they couldn't perform useful work.

Stoddart and his colleagues have now gotten past that hurdle. As they reported in *Science* in December 2021, they immobilized a new breed of molecular pumps on solid particles made from materials known as metal organic frameworks. These particles have a Tinker Toy–like architecture that chemists can control at the atomic level, enabling them to graft their molecular pumps to the particle surfaces in a consistent orientation.

The scientists then showed that by feeding their system a pair of chemicals, they could drive multiple rings onto each grafted rod, increasing their concentration at the surface to a higher level than in solution. Although the researchers haven't done anything useful so far with their minipumps, Stoddart says further tinkering could create teeny machines that pluck carbon dioxide molecules from the air to fight climate change, perhaps by pumping the gas across a membrane that allows it to be captured and sequestered.

Another stride toward making useful molecular pumps came last week, from David Leigh, a chemist at the University of Manchester, and his colleagues. The team immobilized tiny organic molecular rods on micrometer-size plastic beads. Then, like the Stoddart group, they showed that by repeatedly adding a pulse of a chemical fuel, they could thread multiple organic rings onto the rods.

In a twist, the team used two kinds of fluorescent rings, one emitting green light and the other blue, and showed that by delivering pulses of two different chemical fuels, they could thread rings of alternating colors onto the rods, they reported in *Nature Nanotechnology*. One possible use, Leigh says, is a high-density data storage system in which data are written or read by moving rings on or off the rods. Another: using the ring to collect toxins in the blood stream and deliver them to the hollow beads, which could sequester them.

In a final study, published last week in *Nature*, Leigh's team created a rotating motor that spins continuously as long as fuel is present. In this case, a chemical group called a pyrrole-2-carbonyl acts as a rotor that revolves above a stationary group called a phenyl-2-carbonyl. When no fuel is present, another group called a diacid that is attached to the rotor acts as a stop. It bumps into the stationary group, preventing rotation.

A combination of two fuel compounds changes the configuration of the diacid, creating a kind of ratchet. The first compound eliminates the blockage, which allows the rotor to spin; the second locks the rotor and prevents it from spinning backward. Additional pairs of fuel molecules spin it again. "Our motor will spin as long as fuel is present," Leigh says. Although it's not yet clear just what scientists will do with this 26-atom rotary motor, a larger biological analog uses rotational motion to generate adenosine triphosphate, the fuel that powers cells.

For now, the fuel-driven rotor's spin isn't very fast, only about three revolutions per day. But Leigh notes that chemists are still learning the rules for making molecular machines more efficient. The next big hurdle will be finding a way to harness these machines to carry out useful tasks. Minuscule biological motors sustain even the most massive life forms, and Leigh and others think that in industry and medicine, their artificial versions could have an equally potent impact. "It will be a game changer," he says. ■

LOOKING FOR TROUBLE

Trapping bats with Supaporn Wacharapluesadee, who hunts for viruses to understand and prevent pandemic threats

By **Jon Cohen**; *Photography by* **Lauren DeCicca**

Supaporn Wacharapluesadee led a team that in January captured *Rhinolophus* bats (in cloth bags) at Thailand's Khao Ang Rue Nai Wildlife Sanctuary.

In rural Thailand, an elephant sitting in the road is not a charming sight. The massive beasts have a penchant for ripping off bumpers, tusking doors, and sitting on hoods. So in January, when an elephant loomed on the pavement ahead, a van carrying a team of bat researchers on a road 200 kilometers southeast of Bangkok stopped abruptly. As the animal loped toward the van, ears flapping and trunk swinging, the driver slowly backed up. At last, the elephant lumbered back into the other lane and the driver crept past. "That was wild!" a member of the team said.

Its leader, Supaporn Wacharapluesadee, who stands out as mild mannered in a famously mild-mannered culture, fell off her seat laughing with relief. She is used to much smaller, but more consequential, menaces. Within hours, she and her team planned to be in Thailand's Khao Ang Rue Nai Wildlife Sanctuary examining animals for dangerous viruses that might spill over into humans—or already have.

Supaporn is one of the world's most accomplished virus hunters. She is known for her work tracking Nipah virus, a bat-borne pathogen that is less contagious than SARS-CoV-2 but more deadly to humans. She has found bat coronaviruses related to both SARS-CoV, which triggered the epidemic of sudden acute respiratory syndrome (SARS) nearly 2 decades ago, and the virus behind Middle East respiratory syndrome (MERS). And her quest has gained new importance during the COVID-19 pandemic, which likely originated when a bat coronavirus evolved into SARS-CoV-2 and crossed over into humans, perhaps through an intermediate host animal.

She was the first researcher to sequence SARS-CoV-2 outside China—not in an animal, but in an airline passenger—and she is on the trail of its wild relatives. From her base at Chulalongkorn University in Bangkok, Supaporn has made many forays like the one delayed by the elephant. Those outings added precious data points in the hunt for SARS-CoV-2's origin as she identified bat coronaviruses on the virus' family tree—some of which may be its closest relatives yet found.

The 52-year-old scientist's career blossomed over the past decade after she joined PREDICT, a multicountry, well-funded epidemiological program sponsored by the U.S. Agency for International Development (USAID). Until it ended in 2019, the program looked for pathogens in animals and humans to spot new pandemic threats. The World Health Organization in fall 2021 named her a member of its new Scientific Advisory Group for the Origins of Novel Pathogens.

"She's fabulous," says Dennis Carroll, a tropical disease specialist who started PREDICT. "She's demonstrated over the years a really innovative mind in terms of the fieldwork she does, and she's extremely practical, doing really high-quality lab work."

PREDICT's principal investigator, epidemiologist Jonna Mazet of the University of California, Davis, also admires how Supaporn has made her way in a male-dominated field. "She's had to fight for what she got, which is especially impressive in a country like Thailand, where the women are not as supported as they are here in the U.S."

Yet some scientists—including Supaporn's former boss, Thiravat Hemachudha—question whether the type of arduous wild animal surveillance she did on that elephant-interrupted January trip truly makes humans safer. "I don't think it's that valuable, and it may be dangerous," says Thiravat, a neurologist who last year had a complicated falling out with Supaporn that has left her without lab equipment and staff.

Thiravat and other scientists contend that the most efficient way to head off new pandemics is to more aggressively test sick livestock and other animals in contact with people, as well as people with unexplained illnesses, and intensify surveillance of people who often interact with animals harboring dangerous pathogens. "Our motto is: Minimize budget and maximize benefit," Thiravat says.

Supaporn, who hopes to take part in two new viral sleuthing efforts designed to derail spillovers, including a proposed multibillion-dollar Global Virome Project (GVP), says critics are presenting a false choice. To understand viral threats, she says, wildlife surveillance is as important as testing people and livestock. "If we don't do anything, we will not know anything," she says. She and other pathogen hunters say if earlier findings from wild animals had been taken more seriously, "coronavirus" would not have become a common word in every spoken language.

SUPAPORN'S PARENTS MADE fittings for jewelry, and as a child she thought she would become an artist like her brother. But as a teen she realized her talent lay in science. She earned an undergraduate degree in medical technology and spent 10 years working in several diagnostic labs. "When I was young, I was not a communicative person, so working in the lab, there was no need to talk to anyone," Supaporn says. "I thought being a technician was the best job for me."

Snared with butterfly nets and moved to cloth bags, *Rhinolophus* bats are named for their distinctive horseshoe-shaped noses (right).

But when a supervisor hired an outside company to solve an assay problem that she knew how to fix herself, she decided her tech days had ended. "I thought, 'I can do more than that.'"

In graduate school, she studied with Thiravat, who treated people infected with rabies, mainly through dog bites. A related virus that infects Australian bats also causes a rabieslike disease in humans, so she and Thiravat decided in 2002 to start sampling bats in Thailand. The bats carried antibodies to that second virus, indicating its presence in Thailand as well. At the government's behest, the researchers also began sampling bats and other animals for Nipah virus, which emerged in Malaysian pigs and their farmers in 1998, killing up to 75% of infected humans.

Supaporn, Thiravat, and colleagues repeatedly found antibodies to Nipah in *Pteropus*, or flying foxes, the world's largest bats with a 1.5-meter wingspan. Eventually, the team isolated the virus itself from a bat. To dispel folklore about a popular aphrodisiac in Thailand and other Asian countries, they published a paper in *Clinical Infectious Diseases* in 2006 with a startling title: "Drinking Bat Blood May Be Hazardous to Your Health."

To prevent Nipah spillovers, for 2 decades Supaporn has tested humans and pigs in villages near Wat Luang Phrommawat, a 400-year-old temple with a grove of trees where some 10,000 flying foxes roost. She has never found Nipah virus or its immunological footprints in humans or pigs, but Supaporn says the work led the locals to discard fruit that was partially eaten, possibly by the bats.

"I have a responsibility to the community to do education about this risk," she says.

The Nipah studies caught the attention of scientists at the EcoHealth Alliance, a conservation-oriented nonprofit in New York City that was part of PREDICT, and in 2009 it subcontracted with Supaporn to do wildlife surveillance in Thailand. Peter Daszak, who heads EcoHealth, notes that few researchers in the countries where pandemics tend to originate do such work.

"Supaporn's one of those who gets it," says Daszak, who has been scrutinized because of the possibility—dismissed by many scientists as pure speculation—that SARS-CoV-2 leaked from a lab EcoHealth collaborated with at the Wuhan Institute of Virology in China. "And it's not easy for someone to develop their own pathway like she has."

Since then, Supaporn has done multiple studies with EcoHealth and PREDICT. She showed that bat guano used as fertilizer by Thai farmers was contaminated with a coronavirus related to the cause of MERS, and she ranked the spillover potential of different animal viruses. Even before the pandemic, she had described 63 coronavirus sequences detected in 13 species of Thai bats she sampled.

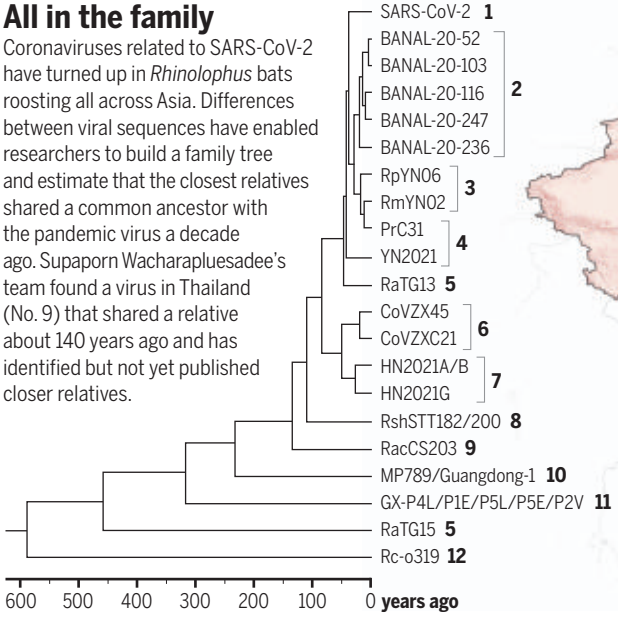
Shi Zhengli, who runs the Wuhan lab and also has come under attack by lab-leak proponents, has collaborated with Supaporn and says they often swap ideas. "Tropical Asia is a hot spot of wildlife-borne emerging infectious diseases," Shi says, "so her job is very important for disease prevention and precaution in the region."

A water pipe at a wildlife sanctuary held a colony of *Rhinolophus* bats. Some sampled in June 2020 harbored a SARS-CoV-2 relative.

PHOTOS: LAUREN DECICCA

All in the family

Coronaviruses related to SARS-CoV-2 have turned up in *Rhinolophus* bats roosting all across Asia. Differences between viral sequences have enabled researchers to build a family tree and estimate that the closest relatives shared a common ancestor with the pandemic virus a decade ago. Supaporn Wacharapluesadee's team found a virus in Thailand (No. 9) that shared a relative about 140 years ago and has identified but not yet published closer relatives.



THE KINDS OF THREATS Supaporn had been tracking became catastrophically real at the beginning of 2020. On 8 January, a passenger arriving from Wuhan at Bangkok's international airport registered hot on thermal scanning equipment. An ear check showed her temperature was 38.1°C. Rome Buathong, a field epidemiologist for the Thai Ministry of Public Health who had set up the scanners 5 days earlier when news arrived about the outbreak in Wuhan, promptly sent the woman to the hospital. All viral tests were negative, so Rome contacted Supaporn, who had worked with him years earlier to screen air passengers for Ebola and Zika viruses.

On 9 January—the day before Chinese researchers first publicly reported SARS-CoV-2's genome—Supaporn discovered the genetic signature of a novel virus in that Wuhan visitor, becoming the first scientist outside China to do so. A database search showed the new virus was closest to a coronavirus in Chinese bats that Daszak and Shi had reported in 2017. "Ten years ago, no one thought bats were important—we thought only about influenza," Rome says. "But Supaporn was very keen to do a lot with bats. Who knew?"

The COVID-19 pandemic in early 2020 halted field research globally, but Supaporn, with funding from the U.S. Department of Defense's Biological Threat Reduction Program, managed that June to send a team to a large cave in western Thailand that is home to a few million bats. The endeavor was part of a general pathogen surveillance effort, but the group hoped to find a clue to SARS-CoV-2's origin by sampling *Rhinolophus* bats, also known as horseshoe bats for the shape of their noses. The genus, comprising more than 100 species, is the main host for SARS-related coronaviruses.

Horseshoe bats live in small colonies that are often hard to find, and the cave didn't yield any. But in a water pipe draining a reservoir that's part of the Khao Ang Rue Nai Wildlife Sanctuary, Supaporn's team trapped 100 *Rhinolophus acuminatus*. Rectal swabs from 13 tested positive for coronaviruses, including one described in *Nature Communications* on 9 February 2021. Dubbed RacCS203, the virus was 91.5% identical in genetic sequence with SARS-CoV-2. That similarity implied a common ancestor from about 140 years ago, according to an analysis led by evolutionary biologists David Robertson and Spyros Lytras of the MRC-University of Glasgow Centre for Virus Research, published online on 8 February in *Genome Biology and Evolution*.

Other researchers found bat coronaviruses related to SARS-CoV-2 in China, Laos, Vietnam, Cambodia, and Japan. One virus from a colony in limestone caves in Laos was 96.8% similar in sequence to the human virus—perhaps a decade removed. Even it is too distant to offer anything more than crumbs on the evolutionary path that led to the pandemic virus. But Robertson is convinced that Asia's bats harbor far closer relatives to SARS-CoV-2. "There's definitely something that's not been sampled," he says.

On the trip this January, Supaporn returned to the sanctuary in search of closer matches. RacCS203, unlike the virus from Laos, does not infect by binding to the human cellular receptor favored by SARS-CoV-2. But antibodies in the blood of bats in the sanctuary powerfully neutralized the pandemic virus, suggesting they may have been infected with a coronavirus that uses that receptor, too.

Some researchers think the bat virus hunt will do little to clarify the pandemic's origin. A distant bat precursor to SARS-CoV-2 might have spread long ago to an intermediate host—perhaps a rat, civet cat, raccoon dog, or pangolin, all known to host bat viruses—and evolved there for years before infecting humans. But Supaporn is betting she'll find revealing clues in bats. "It would be good to fill in the gaps of the origin story in Southeast Asia because in Thailand alone there are 23 *Rhinolophus* species," she says.

Filling in the gaps is a painstakingly slow, expensive, risky, and often hugely unpleasant process. "You're looking for something rare, and you need a ton of samples to pick up the rare thing," Mazet says.

BY THE TIME SUPAPORN'S van passed the elephant and joined the rest of the team at the field site, it was after 4 p.m. With military efficiency, the team—two dozen grad students, ecologists, and veterinarians—set up a lab on the ground floor of an abandoned traditional Thai house on stilts. The first order of business, ironically, was to protect the bats from human viruses, including SARS-CoV-2: Everyone had nasal swabs, which came back negative.

Next, team members put on hairnets, polyethylene coveralls, nitrile gloves, and N95 masks to protect themselves. The temperature was 32°C. Sweat soon soaked every bit of fabric under the zipped-up suits.

CREDITS: (GRAPHIC) K. FRANKLIN/*SCIENCE*; (DATA) DAVID ROBERTSON AND SYPROS LYTRAS/MRC-UNIVERSITY OF GLASGOW CENTRE FOR VIRUS RESEARCH; S. LYTRAS *ET AL*., *GENOME BIOLOGY AND EVOLUTION*, 14, 2 (2022)

A half-dozen men, who also donned rubber boots and mining headlamps, left the lab and went down an adjacent road to the water pipe, home to a few hundred *R. acuminatus*. The group scuttled down a ladder to the pipe's opening. Butterfly nets in hand, they hunched into the tunnel, where the stench of bat feces, urine, and wet fur parked in the nose. A mesh placed over the pipe's opening caught any bats trying to leave the roost.

Emerging from the pipe, the men untangled the mouse-size bats from their nets, a delicate process given the tangle of spiny wings in the mesh and the animals' ice pick teeth. Each bat went into its own cloth bag. Supaporn did not take part in the procedure. "I'm not good at it," she says, noting that she has been bitten several times.

The next day, the team trapped another 50 *Rhinolophus* from the water pipe. The group also captured 50 bats of another species, *Hipposideros*, from beneath a botanical museum at the wildlife sanctuary, so they could be tested to see whether any coronaviruses had jumped from the *Rhinolophus* roosting nearby. After each trapping, Supaporn's team took the animals back to the field station for measurements and tissue samples, aiming to free the bats as quickly as possible to minimize trauma and harm.

"They're one of the more efficient teams I've worked with," says Kevin Olival, an ecologist at EcoHealth. "In many other countries, it would take 5, 6, 7 days to get that many bats."

The field station resembled a production line. At the first cluster of tables, team members weighed each bat, measured head and ear size with a caliper, shone a light through the wing to estimate age from bone joint size, measured wingspan, and tweezed off parasites, saving them in tiny tubes for a separate study. Station two swabbed the anus and mouth, hole punched tissue from a wing, aspirated blood from a capillary, and then brushed red nail polish on toes so no released bat would be sampled twice.

The swabs were later tested for viral genetic material and the wing tissue for DNA confirmation of the species. Supaporn and her collaborators in other countries will test the blood for antibodies against a wide range of paramyxoviruses, influenza viruses, filoviruses, and coronaviruses.

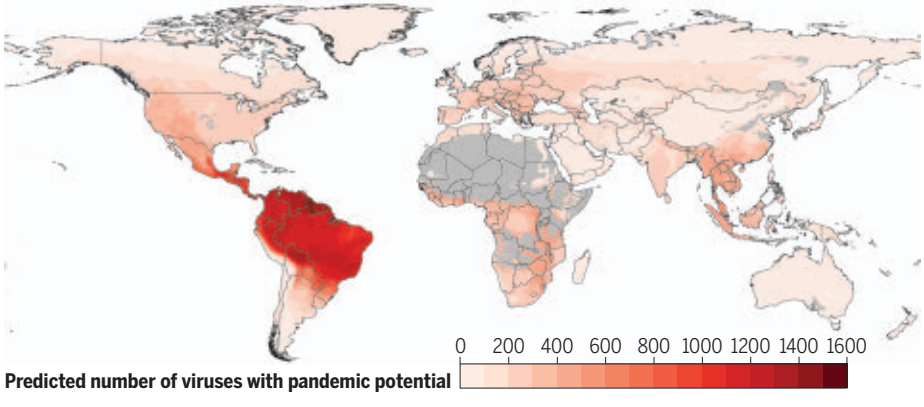
Supaporn worked at a third station, centrifuging bat blood to separate the plasma. She took an occasional breather to pluck a cloth bag from the end of the production line, smiling broadly each time she nudged out a bat with red nail polish and watched it fly off toward home.

Data from the January trapping, to Supaporn's surprise, indicated coronaviruses unrelated to the SARS family in the *Hipposideros* but none in the *Rhinolophus*. Antibody analyses are still underway, and she suspects many *Rhinolophus* will test positive for past infections with SARS-related viruses.

Another foray, a March 2021 expedition to a cave west of Bangkok, yielded two new SARS-CoV-2–related coronaviruses in a species of *Rhinolophus* called *R. pusillus*. Supaporn analyzed them with Linfa Wang, a specialist in emerging infectious diseases

Hot zones

Researchers proposing a Global Virome Project have mapped regions where unknown viruses in wild mammals are most likely to spark human pandemics. Their predictions draw on data on known viruses, traits that predispose viruses to infecting humans, and human populations. Although Southeast Asia has been a hot spot of outbreaks driven by bat viruses, the Amazon region's pandemic potential may be much higher.



CREDITS: (PHOTO) LAUREN DECICCA; (GRAPHIC) K. FRANKLIN/*SCIENCE*; (DATA) ECOHEALTH ALLIANCE; AFTER OLIVAL *ET AL.*, *NATURE* 2017

Swarmed by moths, Supaporn Wacharapluesadee's team worked into the evening sampling blood and other bat tissues at a makeshift lab (left). A hole punch in the wing will provide DNA to identify the species of this *Hipposideros* bat.

PHOTO: LAUREN DECICCA

at Duke-NUS Medical School in Singapore who in 2013 co-wrote a paper with Shi and Daszak describing the first bat coronavirus linked to SARS-CoV. Wang says he and Supaporn plan to report that in some parts of the viral surface protein that docks onto animal cells, the new viruses "have a closer relationship with SARS-CoV-2 than any other previously found in bats."

HEROIC AS SUCH WILDLIFE surveillance may seem, some scientists question its value for heading off future pandemics.

PREDICT, which received $207 million from USAID from 2009 to 2019, discovered 959 novel viruses and identified hot spots for spillovers to humans, along with training Supaporn and nearly 7000 other researchers. "We were building their surveillance systems with them," Mazet says.

Edward Holmes, an evolutionary biologist at the University of Sydney, applauds PREDICT's training efforts but has doubts about whether the effort made the world safer. "It produced a fair amount of sequence data, but has it actually predicted anything?" he asks. "I don't really know. It didn't get SARS-CoV-2."

Carroll, who retired from USAID in 2019, and scientists who participated in PREDICT contend that the project clarified what drives spillovers, such as the wildlife trade at markets and deforestation. PREDICT's supporters also say it pinpointed sites where outbreaks are most likely. But Carroll readily acknowledges PREDICT's limitations. "Its scope was too small to have a meaningful impact," he says.

Carroll, Mazet, Daszak, and a small group of other researchers see PREDICT as a trial run for a much bigger effort: a GVP that aims to identify 75% of the viruses most likely to spill over within 10 years, at an estimated cost of $4 billion. GVP organizers, who started to flesh out the idea 6 years ago, had hoped to launch in 2020 with support from China and Thailand. The pandemic derailed their plans—but also underscored the need, Carroll, Supaporn, and other researchers argued last year in a commentary in *The BMJ*.

Holmes has assailed the idea of the GVP since it was first floated. "It's absolute nonsense," he says. "It's too big a bloody arena." Nearly all threatening pathogens are RNA viruses, which mutate at a fast clip, constantly creating new variants, Holmes notes. "You've got an amazing diversity of viruses that are continually turning over, so how would you then decide, 'That's the one that I'm worried about?'" he asks. "Surveillance is infinitely better and more cost-effectively directed at humans."

Supaporn counters that the goal of wildlife surveillance isn't to characterize every potential viral threat, but rather to learn how viruses evolve. And she is convinced that this work can predict the most likely future pathogens. "Even a general sense of this is extremely valuable to public health planning efforts," she says. "Learn, understand, prepare."

THOSE ARGUMENTS MAY HAVE contributed to Supaporn's falling out with Thiravat, which forced her to walk away from the institution he heads, the Health Science Centre of the Thai Red Cross Emerging Infectious Diseases program at King Chulalongkorn Memorial Hospital. She is now at its sister Clinical Centre, without her equipment and trained technicians. Work like Supaporn's promises more risks than benefits, Thiravat contends. "Wildlife surveillance may introduce human pathogens to wildlife and vice versa." As for SARS-CoV-2, he believes it was not a natural jump of a virus from animals to humans. "It was a product of lab leak of virus after manipulation," he asserts. (Thiravat has also advocated using the antiparasitic drug ivermectin to treat COVID-19, even though multiple studies have shown it is ineffective.)

Thiravat contends that Supaporn siphoned off about $400,000 from grants. But an investigation conducted by the Thai Red Cross Society exonerated her in July 2021, concluding in a letter (which she supplied to *Science*) that there was "no evidence of financial conduct contrary to [her employer's] regulations."

Some Supaporn supporters say Thiravat is jealous of the attention she has received for her coronavirus work during the pandemic. She says the problem began when she challenged things he said to his supervisors, which she did not want to discuss in detail. "I've always respected him—he is my mentor and an intelligent clinician and scientist," she says. "And I'm lucky that even though I have some politics in the lab, people outside Thailand don't think that I'm wrong, and they support me."

Supaporn's setbacks mean she must now rely on colleagues, including Wang, to complete the lab analyses of samples her team collects in the field. But she's upbeat about her future.

In early March, she met with researchers from a new $125 million, 5-year project launched last year by USAID called Discovery & Exploration of Emerging Pathogens—Viral Zoonoses (DEEP VZN), taking them to the flying fox colony in the trees at Wat Luang Phrommawat. While she waits to see whether DEEP VZN makes her a collaborator and whether the GVP finds funding, Supaporn has enough grant money to continue her fieldwork for the time being.

For now, she focuses on training students and embracing the many unknowns she faces. "It's a Buddhist teaching," she says. "Uncertainty is certainty."

Which could also be a motto for the entire pandemic prevention enterprise. ■

This story was supported by a grant from the Alfred P. Sloan Foundation.

INSIGHTS

PERSPECTIVES

CRYSTALLOGRAPHY

The fascinating world of biogenic crystals

The diverse properties of these crystals may lead to a variety of applications

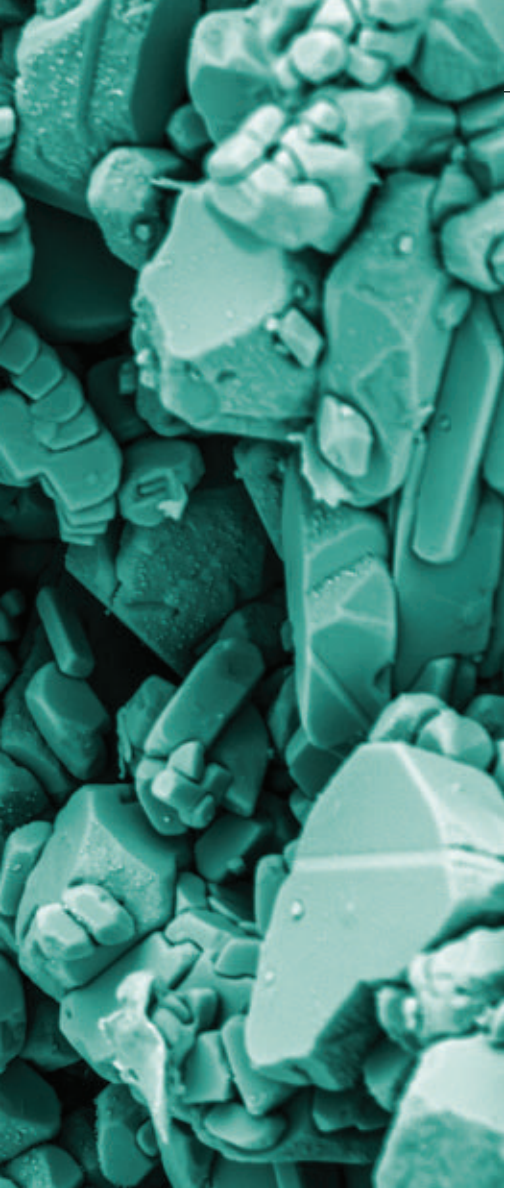
By **Jolanta Prywer**

Biogenic crystals are crystals that grow inside or under the influence of living organisms. These crystals are incredibly diverse, and the process of their growth is extremely complex. Despite their ubiquity, only several dozen biogenic crystals have been identified and studied. On page 312 of this issue, Avrahami *et al.* (*1*) present studies of coccoliths, which are micrometer-sized calcite ($CaCO_3$) single plates formed by coccolithophores, a type of single-celled algae. The coccolithophores form around themselves a calcite shell called a coccosphere, which is composed of dozens of coccolith plates. Before discussing the most important and specific aspects of this study, it is worth exploring the diverse world of biogenic crystals.

The habits of biogenic crystals—defined as the characteristic external appearances of the crystal—are incredibly diverse. They range from simple regular prismatic crystals through spherical aggregates to individual needles with sharp points. In the case of calcium oxalate [$CaC_2O_4 \cdot (H_2O)_x$] crystals found in over 200 plant families (*2*), their habits and their presence in specific plant tissues determine their functions. For example, needle-like crystals found in leaves may play a defensive role against herbivores (*3*). Sometimes these needles have grooves that can help transport toxins to the nibbled parts of the plant. Biogenic crystals can also be used to help manage nutrients and growth. For example, calcium oxalate crystals, found in plants, are thought to serve as a calcium reservoir for maintaining a proper ionic balance (*4*). The same crystals may also be used to help control the level of solubility for potentially toxic substances such as oxalic acid.

There also exist other, perhaps more peculiar, uses for biogenic crystals. For example, magnetotactic bacteria produce nanometer-sized magnetic crystals in their intracellular membrane vesicles (*5*) to help orient themselves using Earth's magnetic field and navigate toward certain aquatic environments that favor their survival (*6*).

IMAGE: NANO CREATIVE/SCIENCE SOURCE



Biogenic crystals, such as coccoliths (pictured here) formed by single-celled algae, play an essential role in many biological processes.

The magnetic properties of these crystals have been studied for their potential applications—for instance, enabling precision drug delivery in cancer treatments (*7*). Another distinctive example is the guanine nanocrystals found in the skin of chameleons, which play a role in the color change ability of chameleons (*8*). In addition, the deeper skin cells contain slightly larger guanine crystals that can reflect sunlight in the infrared range, which means that guanine crystals also provide thermal protection for the chameleons. This, like the other biogenic crystals mentioned, has also provoked a search for possible applications, such as a chameleon-inspired electronic skin with touch-controlled color change (*9*). Biogenic crystals can also be found in some pathological processes related to human health. An example is the struvite crystal, which is formed in the urinary tract when infected by urease-positive bacteria. Bacteria actively participate in the process of growing these crystals, such as by affecting their porosity and the characteristic surface structure (*10*, *11*). Through these crystals, bacteria can enhance their adhesion to the epithelium of urinary tract, making them more difficult to be dislodged by the urine stream (*12*). Therefore, the study of biogenic crystals is also of interest to the medical community.

Adding to the diverse roster of biogenic crystals, Avrahami *et al*. dive into microscopic detail of coccoliths grown by the marine algae *Calcidiscus leptoporus*. Coccoliths are micrometer-sized plates of calcite crystals grown around an organic substrate known as the base plate. Researchers have created models to explain how organisms crystallize such complex structures. One such popular model is the V/R model, which prescribes the coccoliths to have two different crystal units: a radial (R) unit with the crystallographic *c* axis oriented parallel to the coccolith plane, and a vertical (V) unit with the *c* axis perpendicular to the coccolith plane (*13*).

After examining the stages of coccolith formation using various three-dimensional imaging techniques, Avrahami *et al*. propose that the orientations of the R and V units are determined by having their different edges attached to the base plate. In this model, R and V units are one and the same rhombohedron-shape calcite crystal. Rhombohedra placed on their acute edges are R units, and those on obtuse edges are V units. Through careful characterization of the morphology and orientation of calcite crystals at various stages of growth, the authors conclude that these rhombohedron-shape calcite crystals are built of only one facet set—the {104}, which contains six symmetry-related facets.

Avrahami *et al*. show that the facets of the {104} calcite rhombohedron found in coccoliths can have different growth rates, breaking the symmetry of the rhombohedron. The structure of the coccolith can be explained by the anisotropic growth rates of the {104} rhombohedron facets, which depend on their orientation in relation to the concentration gradient of calcium and carbonate ions. A similar symmetry-breaking phenomenon has also been observed, for example, in the case of the magnetite crystals from the aforementioned magnetotactic bacteria (*14*). Such anisotropy of the growth rates may occur due to an anisotropy in the environment or of the growth sites, which may be the result from an uneven flux of ions passing through the intracellular membrane surrounding the crystal.

If the idea of symmetry breaking and anisotropy of the growth rates of symmetrically equivalent facets is correct, then other interesting phenomena can be expected—for example, an increase in the size of fast-growing facets. Growing crystals adopt a morphology such that the bounding facets of the crystal have a low surface energy, which corresponds to the slow-growing facets, whereas the fast-growing facets tend to disappear because of a higher surface energy. By looking at the {104} calcite rhombohedron as seen in coccoliths, one can see how the growth anisotropy can alter the overall appearance of a crystal (*1*). Although it is little wonder that the slow-growing facets can grow larger, the geometry of the {104} rhombohedron is such that the fast-growing facets can grow larger as well. The phenomenon of increasing the size of fast-growing facets, rather than their disappearance, is of great interest but not very often observed and is related to the geometry of the crystal (*15*).

In the process of biogenic precipitation of calcite in the form of coccoliths, coccolithophores release carbon dioxide, and simultaneously, in the process of photosynthesis, they capture carbon dioxide. This process takes place on an enormous scale in the oceans and is one of the basic factors regulating the carbon dioxide (CO_2) and carbon cycle in nature. The difference between capture and excretion of CO_2 by coccolithophores is an important consideration for modeling carbon cycling in the oceans and therefore has relevance for climate change discussions. The research presented by Avrahami *et al*. is thus in line with the current research trends and can help to assess the role of coccolithophore species *C. leptoporus* with respect to the global carbon cycle. ■

Institute of Physics, Lodz University of Technology, ul. Wólczańska 217/221, 93-005 Łódź, Poland Email: jolanta.prywer@p.lodz.pl

REFERENCES

1. E. M. Avrahami, L. Houben, L. Aram, A. Gal, *Science* **376**, 312 (2022).
2. P. A. Nakata, *Plant Sci.* **164**, 901 (2003).
3. M. L. Peschiutta, S. J. Bucci, G. Goldstein, F. G. Scholz, *Arthropod-Plant Interact.* **14**, 727 (2020).
4. P. V. Monje, E. J. Baran, *Plant Physiol.* **128**, 707 (2002).
5. D. A. Bazylinski, R. B. Frankel, *Nat. Rev. Microbiol.* **2**, 217 (2004).
6. R. E. Kopp, J. L. Kirschvink, *Earth Sci. Rev.* **86**, 42 (2008).
7. A. Basit, J. Wang, F. Guo, W. Niu, W. Jiang, *Microb. Cell Fact.* **19**, 197 (2020).
8. J. Teyssier, S. V. Saenko, D. van der Marel, M. C. Milinkovitch, *Nat. Commun.* **6**, 6368 (2015).
9. H.-H. Chou *et al*., *Nat. Commun.* **6**, 8011 (2015).
10. H. Li, Q.-Z. Yao, Y.-Y. Wang, Y.-L. Li, G.-T. Zhou, *Sci. Rep.* **5**, 7718 (2015).
11. Y. Rui, C. Qian, *J. Cryst. Growth* **570**, 126214 (2021).
12. J. Prywer, A. Torzewska, *Sci. Rep.* **9**, 17061 (2019).
13. F. C. Meldrum, B. R. Heywood, S. D. Mann, R. B. Frankel, D. A. Bazylinski, *Proc. Biol. Sci.* **251**, 237 (1993).
14. J. R. Young, J. M. Didymus, P. R. Brown, B. Prins, S. Mann, *Nature* **356**, 516 (1992).
15. J. Prywer, *Prog. Cryst. Growth Charact. Mater.* **50**, 1 (2005).

10.1126/science.abo2781

CATALYSIS

Fullerenes make copper catalysis better

Ethylene glycol can be reliably produced by mild hydrogenation of dimethyl oxalate

By **Edmond Gravel** *and* **Eric Doris**

Catalysis plays a central role in modern synthetic chemistry by reducing the activation energy needed for a reaction to take place. Among metal catalysts, copper (Cu) is an inexpensive and highly versatile catalyst with applications (*1*) such as in hydrogenation reactions used in the bulk production of ethylene glycol from dimethyl oxalate (DMO). During the DMO-to-ethylene glycol transformation, high pressure of hydrogen gas (H_2) is usually needed. This can be problematic from an engineering and safety perspective, and can also lead to catalyst deactivation. On page 288 of this issue, Zheng *et al.* (*2*) provide a solution to the pressure critical point by associating a conventional Cu catalyst with fullerenes, which are molecules made of carbon atoms organized in a spherical structure such as in C_{60}. The method stabilizes the catalytic species and enables the hydrogenation of DMO into ethylene glycol under mild conditions of pressure.

Université Paris-Saclay, CEA, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE), Département Médicaments et Technologies pour la Santé (DMTS), Service de Chimie Bioorganique et de Marquage (SCBM), 91191 Gif-sur-Yvette, France. Email: eric.doris@cea.fr

Ethylene glycol is a commodity chemical with applications in many areas of everyday life. It is used as a solvent, a coolant in automotive radiators, a deicing fluid for aircraft, and a building block in the synthesis of polyester fibers and resins (*3*). The current global ethylene glycol production capacity is approximately 42 million metric tons per year and is forecast to exceed 70 million metric tons by 2025 (*4*). The industrial synthesis of ethylene glycol relies mostly on the oxidation of ethylene followed by thermal hydration of the resulting ethylene oxide (see the figure), a process established in the 1930s by the Union Carbide Corporation. However, ethylene is a petroleum-based chemical, and other routes have recently emerged from more sustainable feedstocks. The path from synthesis gas (syngas), a gasification product of coal or biomass (*5*) that yields primarily a mixture of hydrogen and carbon monoxide (CO), is of particular interest. This two-step process involves converting CO into DMO by means of oxidative coupling with methanol, followed by hydrogenation of DMO (*6*).

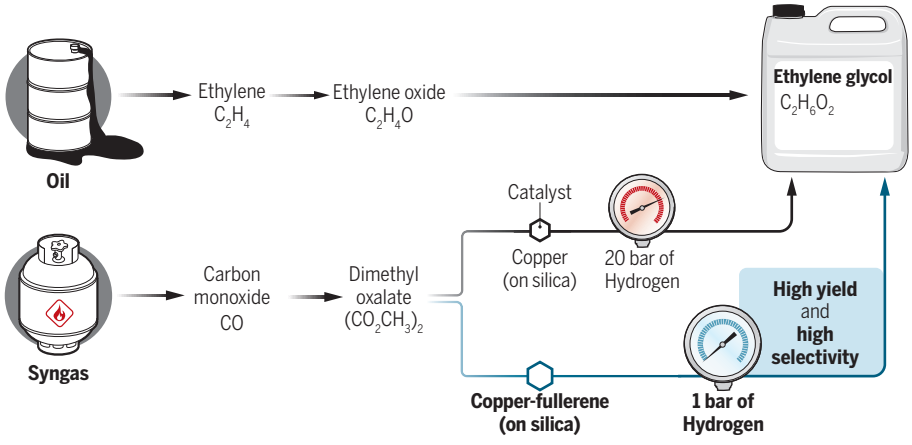
The most challenging part of this synthetic sequence is the hydrogenation of DMO to ethylene glycol because high pressures of hydrogen gas at elevated temperatures are required for the reaction to be effective (*7*). Metals such as Cu catalyze the hydrogenation of DMO (*8*), but the drastic conditions required for the reaction to occur can lead to catalyst deactivation (*9*) as well as safety and environmental issues.

To stabilize Cu and make the catalytic hydrogenation of DMO more robust, Zheng *et al.* came up with the idea of associating Cu to an electron-buffer fullerene unit to keep the right balance between the catalytically active forms of the metal throughout the hydrogenation process. Cu can exist in various oxidation states, ranging from the elemental Cu^0 to the ion Cu^{4+}, but only Cu^0 and Cu^+ are useful for the hydrogenation reaction of DMO to ethylene glycol. It is thought that Cu^0 and Cu^+ act in a synergistic fashion, with Cu^0 promoting the dissociation of H_2, and Cu^+ promoting the addition of H to DMO (*10*). Fullerenes, specifically the C_{60} "buckyballs" (*11*), are electronically active and can accept and, to some extent, give back electrons. This electron-buffering property was exploited by Zheng *et al.* to design a hybrid catalyst that associates Cu and C_{60} on an inert silica platform: C_{60}–Cu/SiO_2. In this setup, the fullerenes electronically interact with Cu to protect the unstable Cu^+ from oxidation and reduction and maintain a catalytically active Cu^0/Cu^+ ratio. Evidence of the buffering effect of fullerenes on Cu was obtained through various techniques, including cyclic voltammetry, and further supported with theoretical calculations. Zheng *et al.* demonstrated that C_{60} could act sequentially as a single-electron acceptor (from Cu^0) and a donor (to Cu^{2+}) to prevent the active Cu^+ catalyst payload from changing oxidation state.

When compared with a conventional Cu/SiO_2 catalyst, the fullerene-buffered Cu catalyst C_{60}–Cu/SiO_2 showed strong performances in the vapor phase conversion of DMO. And in the case of C_{60}–Cu/SiO_2, the transformation could be carried out under ambient pressure of H_2 gas, whereas until now, standard catalysts required high hydrogen pressures of more than 20 bar. At ambient pressure, the C_{60}–Cu/SiO_2 catalyst leads to a 10-fold increase in ethylene glycol yield (98%) compared with that of Cu/SiO_2 and is more selective because it generates fewer by-products. The use of 12 g of C_{60}–Cu/SiO_2 permitted the flow production of ethylene glycol on a multikilogram scale with no alteration of the catalyst over 1000 hours of operating time at ~180°C. After the reaction was complete, the catalyst could

Toward a more sustainable ethylene glycol production

Ethylene glycol is classically produced from petroleum-based ethylene (top). A more sustainable production method uses dimethyl oxalate obtained through oxidative dimerization of carbon monoxide from syngas but requires a hydrogen pressure of >20 bar (middle). Zhang *et al.* designed a copper-fullerene catalyst that eliminates the pressure requirement (bottom).



GRAPHIC: N. CARY/*SCIENCE*

be recovered and reused, with no loss in activity. C_{60}–Cu/SiO_2 was further used in the hydrogenation of DMO-related substrates, again outperforming Cu/SiO_2. As a test, ethyl acetate was efficiently reduced to ethanol by means of ambient pressure hydrogenation over C_{60}–Cu/SiO_2, whereas Cu/SiO_2 failed to catalyze any transformation under the same reaction conditions.

The development of reliable and durable catalysts is of prime importance to the chemical industry. The concept of associating fullerenes to Cu provides a distinctive answer to in situ stabilization of the active catalytic species. This buffering effect of C_{60} resulted in a particularly efficient system that was valorized in the ambient pressure hydrogenation of challenging substrates but also in the electrochemical reduction of CO_2 into CO. Because DMO can be synthesized from CO, C_{60}–Cu/SiO_2 can be seen as a two-in-one catalyst to produce raw CO—the precursor of DMO—and the subsequent reduction of DMO to ethylene glycol.

"The method stabilizes the catalytic species and enables the hydrogenation of DMO into ethylene glycol under mild conditions of pressure."

Considering the commercial availability of fullerenes, the hydrogenation technology using C_{60}–Cu/SiO_2 is expected to be soon sufficiently mature, pointing to its industrialization in a not-too-distant future for the cost-effective and sustainable production of ethylene glycol. ■

REFERENCES AND NOTES

1. A. Gopinathan, S. Saranya, Eds., *Copper Catalysis in Organic Synthesis* (John Wiley & Sons, 2020).
2. J. Zheng *et al.*, *Science* **376**, 288 (2022).
3. H. Yue, Y. Zhao, X. Ma, J. Gong, *Chem. Soc. Rev.* **41**, 4218 (2012).
4. https://www.statista.com/statistics/1067418/global-ethylene-glycol-production-capacity
5. J. Ren, J.-P. Cao, X.-Y. Zhao, F.-L. Yang, X.-Y. Wei, *Renew. Sust. Energ. Rev.* **116**, 109426 (2019).
6. H. Song *et al.*, *Chin. J. Catal.* **34**, 1035 (2013).
7. R.-P. Ye *et al.*, *ACS Catal.* **10**, 4465 (2020).
8. Z. He, H. Lin, P. He, Y. Yuan, *J. Catal.* **277**, 54 (2011).
9. Y. Zhao *et al.*, *Ind. Eng. Chem. Res.* **59**, 12381 (2020).
10. Y. Sun, F. Meng, Q. Ge, J. Sun, *ChemistryOpen* **7**, 969 (2018).
11. H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, R. E. Smalley, *Nature* **318**, 162 (1985).

ACKNOWLEDGMENTS

The Service de Chimie Bioorganique et de Marquage (SCBM) is a partner of NOMATEN, a Centre of Excellence in Multifunctional Materials for Industrial and Medical Applications. This paper is dedicated to the memory of our former Head of Department, Dr. Bernard Rousseau.

10.1126/science.abo3155

IMMUNOLOGY

Regulatory CD8⁺ T cells suppress disease

A subset of CD8⁺ T cells regulate chronic inflammation by killing pathogenic CD4⁺ T cells

By **Anaïs Levescot** *and* **Nadine Cerf-Bensussan**

In vertebrates, precise orchestration of immune responses largely relies on two major subsets of T lymphocytes. CD8⁺ T cells are cytotoxic and can eliminate virus-infected host cells, whereas CD4⁺ T cells provide signals that help the activation of CD8⁺ T cells and antibody production by B lymphocytes. To avoid hyperactivation that may lead to irreparable tissue damage but also to prevent inadvertent responses against host tissues, a complex set of regulatory mechanisms exist. CD4⁺ regulatory T cells (T_{regs}) have been well characterized (*1*). By contrast, the existence of regulatory CD8⁺ T cells is a matter of debate. In mice, regulatory CD8⁺ T cells elicited in a model of multiple sclerosis reduced disease severity by suppressing autoimmune CD4⁺ T cells (*2*). On page 265 of this issue, Li *et al.* (*3*) provide evidence that a subset of human CD8⁺ T cells also selectively suppress pathogenic CD4⁺ T cells during autoimmune or infectious diseases, including flu and COVID-19.

A hallmark of the CD8⁺ suppressor T cells in mice is the surface expression of the Ly49F receptor (*3*, *4*). Li *et al.* found that a subset of human CD8⁺ T cells express functional homologs of Ly49F, the inhibitory killer cell immunoglobulin-like receptors (KIRs) KIR3DL1 and KIR2DL3. KIRs regulate the killing activity of natural killer cells, a subset of cytotoxic lymphocytes distinct from CD8⁺ T cells. They found that CD8⁺ T cells expressing KIR3DL1 and/or KIR2DL3 (KIR⁺CD8⁺ T cells) represented only 0 to 3% of blood CD8⁺ T cells in healthy individuals but increased significantly in the blood and inflamed tissues of patients with diverse autoimmune conditions, as well as in severe cases of COVID-19 with evidence of vasculitis.

To demonstrate the regulatory role of human KIR⁺CD8⁺ T cells, the authors investigated celiac disease (CeD). This common intestinal inflammatory disease shares many features with autoimmune diseases but is induced in genetically predisposed individuals by gluten, a major component of the human diet. The central pathogenic role of intestinal CD4⁺ T cells that recognize a well-defined set of deamidated gluten peptides is well established (*5*). After showing that KIR⁺CD8⁺ T cells expanded in the blood and in intestinal tissues of CeD patients exposed to gluten, Li *et al.* demonstrated that KIR⁺ but not KIR⁻ CD8⁺ T cells could block the expansion of gluten-specific CD4⁺ T cells in vitro. Supporting the view that, like their putative counterpart in mice, human KIR⁺CD8⁺ T cells act through killing, they displayed a transcriptomic profile of effector cytotoxic cells and required direct contact, which was associated with CD4⁺ T cell apoptosis.

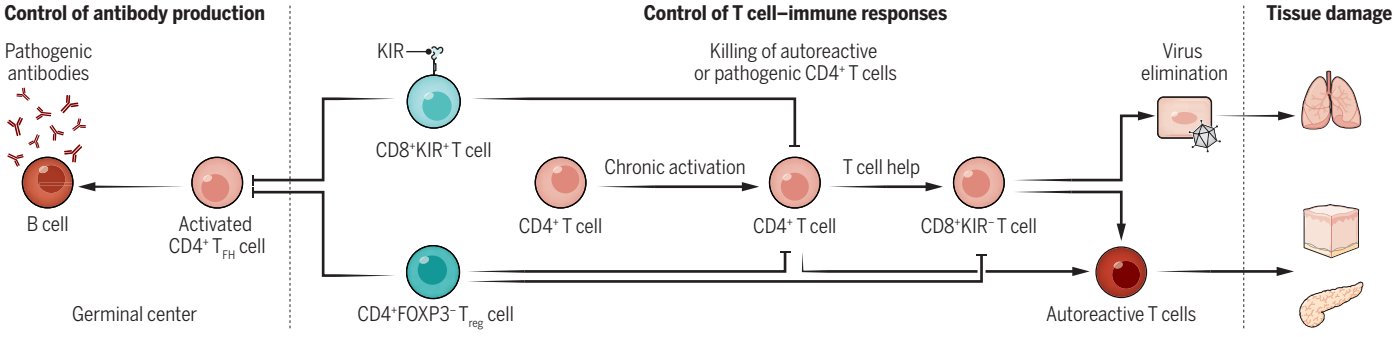
Li *et al.* next searched for evidence that human KIR⁺CD8⁺ T cells represent a distinct cell lineage. They showed that the transcriptomic profile and the T cell receptor (TCR) repertoire of KIR⁺CD8⁺ T cells are distinct from those of KIR⁻CD8⁺ T cells but show similiarities across cells from healthy controls and patients with diverse pathological conditions, such as autoimmune diseases or infections by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and influenza virus. Of 800 genes differentially expressed between the two subsets, ~100 genes were previously found upregulated in mouse Ly49F⁺CD8⁺ T cells and may represent a shared signature. Moreover, one of these genes encodes the transcription factor Helios, which is indispensable to maintain the suppressive functions of forkhead box P3 (FOXP3)⁺CD4⁺ T_{reg} cells and Ly49F⁺CD8⁺ T cells in mice (*6*). Yet, the gene expression profile of KIR⁺CD8⁺ T cells is not clearly distinguishable from that of terminally differentiated cytotoxic T cells.

To further demonstrate that KIR⁺CD8⁺ T cells are bona fide regulatory T cells, Li *et al.* engineered a mouse model selectively depleted of Ly49F⁺CD8⁺ T cells. These mice did not show spontaneous autoimmunity. Yet, upon infection by lymphocytic choriomeningitis virus (LCMV), they developed a

Université de Paris Cité, Imagine Institute, Laboratory of Intestinal Immunity, INSERM UMR 1163, Paris, France. Email: nadine.cerf-bensussan@inserm.fr

Regulatory CD8⁺ T cells control pathogenic immune responses

During immune responses against pathogens and viruses, activated CD4⁺ T cells provide help to B cells that produce neutralizing antibodies and to CD8⁺ T cells that kill infected host cells. Chronic activation of immune responses can cause tissue damage. Human CD8⁺ T cells expressing the inhibitory killer cell immunoglobulin-like receptors (KIRs) KIR3DL1 and/or KIR2DL3 expand during chronic inflammation and kill activated pathogenic CD4⁺ T cells and T follicular helper (T_{FH}) cells, complementing the regulatory functions of forkhead box P3 (FOXP3)⁺ CD4⁺ T_{reg} cells.



lupus-like autoimmune disease with massive expansion of CD4⁺ T follicular helper (T_{FH}) cells and germinal centers despite normal virus clearance. This is consistent with previous findings that mouse Ly49F⁺CD8⁺ T cells can kill activated T_{FH} cells and thereby prevent excessive activation of germinal centers and production of pathogenic antibodies (*4*). It also suggests, as proposed by Li *et al.*, that CD8⁺ regulatory T cells are not involved in suppressing pathogenic immune responses in viral infections (e.g., COVID-19 and influenza) but rather limit the damage induced by autoreactive T cells that are cross-reactive to foreign antigens (*3*, *7*). Overall, Li *et al.* show that a subset of CD8⁺ T cells expressing Ly49F in mice and KIR in humans can exert feedback control on autoimmune and/or pathogenic CD4⁺ T cells (see the figure). However, many gaps in our understanding remain.

It is unclear what is the nature of the signal or signals that induce the activation of regulatory KIR⁺CD8⁺ T cells and enable the selective killling of antigen-specific CD4⁺ T cells. In mice, the regulatory function of Ly49F⁺CD8⁺ T cells is triggered by engagement of their TCR by endogenous peptides presented by either classical or nonclassical class I major histocompatibility complexes (MHCs) (*2*, *4*). This may also be true for human KIR⁺CD8⁺ T cells and implies that the endogenous antigens recognized by KIR⁺CD8⁺ T cells are selectively expressed by chronically activated CD4⁺ T cells. Indeed, recognition of self-antigens by CD8⁺ T cells is a major mechanism driving tissue destruction in autoimmune diseases. In a mouse model of type I diabetes, disease-causing CD8⁺ T cells that killed pancreatic cells were recently identified. Comparing transcriptomes and TCR repertoires of autoimmune CD8⁺ T cells and KIR⁺CD8⁺ T cells may help to understand whether their opposite functions derive from distinct differentiation programs or from their recognition of endogenous peptides expressed by different cellular targets (*8*).

Evidence that the suppressor function of Ly49F⁺CD8⁺ T cells can be activated by peptides expressed by pathogenic CD4⁺ T cells was obtained in a mouse model of multiple sclerosis. In this model, driven by pathogenic CD4⁺ T cells directed against a myelin-specific autoantigen, MHC class I–restricted peptides that selectively bound the TCRs of the protective Ly49F⁺CD8⁺ T cells were identified (*2*). Moreover, immunization of mice with these peptides elicited Ly49F⁺CD8⁺ T cells that protected against disease and killed the pathogenic myelin-specific CD4⁺ T cells. This effect was disease specific because no protection was observed against autoimmune uveitis induced by an eye-specific autoantigen (*2*). It will be interesting to extend these findings to humans and to define whether protection by KIR⁺CD8⁺ T cells is disease specific.

The mechanisms that drive the differentiation, expansion, and activation of KIR⁺CD8⁺ T cells are also important unknowns. A hallmark of these cells is their expression of KIR. This seems to be the consequence of their chronic activation, although the exact mechanism remains to be elucidated. Studies in mice suggest that KIR could promote their survival (*9*) but also attenuate their regulatory function (*10*). Another important feature of KIR⁺CD8⁺ T cells is their expression of Helios. This transcription factor maintains the regulatory function of Ly49F⁺CD8⁺ T cells, at least partly through its role in promoting transcription downstream of signal transducer and activator of transcription 5 (STAT5) (*6*). This is consistent with the established role of interleukin-15 (IL-15), a cytokine that signals mainly through STAT5, in the survival of Ly49F⁺CD8⁺ T cells in vivo (*4*). However, this further blurs the distinction from other cytotoxic CD8⁺ T cells that are also dependent on IL-15 for survival (*11*). This is notably the case in CeD, in which CD8⁺ intraepithelial lymphocytes activated through cooperative interactions between CD4⁺ T cells and IL-15 are thought to play a major role in tissue destruction (*12*, *13*). There is currently much interest in the therapeutic modulation of IL-15 to enhance antitumor responses or, conversely, to inhibit autoimmune tissue destruction. If both effector and regulatory CD8⁺ T cells rely on IL-15 for survival and activation, this strategy should be considered carefully.

Overall, Li *et al.* unveil an intriguing population of human regulatory CD8⁺ T cells that can suppress inflammatory responses through killing activated pathogenic T cells. Targeting KIR⁺CD8⁺ T cells might emerge as a new strategy to suppress undesirable self-reactivity in autoimmune or infectious diseases. ■

REFERENCES AND NOTES

1. J. B. Wing, A. Tanaka, S. Sakaguchi, *Immunity* **50**, 302 (2019).
2. N. Saligrama *et al.*, *Nature* **572**, 481 (2019).
3. J. Li *et al.*, *Science* **376**, eabi9591 (2022).
4. H. Nakagawa, L. Wang, H. Cantor, H.-J. Kim, *Adv. Immunol.* **140**, 1 (2018).
5. B. Jabri, L. M. Sollid, *J. Immunol.* **198**, 3005 (2017).
6. H.-J. Kim *et al.*, *Science* **350**, 334 (2015).
7. L. F. Su, B. A. Kidd, A. Han, J. J. Kotzin, M. M. Davis, *Immunity* **38**, 373 (2013).
8. S. V. Gearty *et al.*, *Nature* **602**, 156 (2022).
9. S. Ugolini *et al.*, *Nat. Immunol.* **2**, 430 (2001).
10. L. Lu, H.-J. Kim, M. B. F. Werneck, H. Cantor, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19420 (2008).
11. J. C. Nolz, M. J. Richer, *Mol. Immunol.* **117**, 180 (2020).
12. V. Abadie *et al.*, *Nature* **578**, 600 (2020).
13. N. Korneychuk *et al.*, *Gastroenterology* **146**, 1017 (2014).

10.1126/science.abp8243

GRAPHIC: A. MASTIN/*SCIENCE*

SOFT ROBOTICS

Strong and fast hydrogel actuators

Plant cells inspire a hydrogel actuator that achieves ultrastrong and fast actuation

By **Zhen Jiang**[1] and **Pingan Song**[1,2]

The movements of soft-bodied animals have long inspired scientists to design soft actuators (*1*) that can convert various forms of energy into mechanical work. Hydrogels (*2*) hold the potential to close the performance gap between synthetic actuators and biological organisms because of their similarity to soft tissues, excellent biocompatibility, and large deformations. Expansion toward soft robots and artificial muscles challenges their status, calling for hydrogels with large actuation forces and fast responses to external stimuli. However, existing hydrogel actuators usually exhibit low actuation forces (≤2 N) and slow responses. On page 301 of this issue, Na *et al.* (*3*) report bypassing state-of-the-art hydrogel actuators to achieve an ultrahigh actuation force (730 N) and high speed by combining turgor design and electroosmosis.

Generally, a hydrogel actuator works through a change of osmotic pressure in the network. The resulting pressure, up to a few megapascals, cannot be fully harnessed as actuation forces because it is balanced with the elastic restoring stress of the network upon a swelling equilibrium. One attractive method is the use of dissipation mechanisms—e.g., dual crosslinking (*4*) and double networks (*5*)—to improve the mechanical strength of hydrogels, leading to enhanced actuation forces. This mechanism does not contribute to the actuation speed, which largely depends on the hydrogel porosity.

Compared with osmotic mechanisms, nonosmotic mechanisms can be used to create hydrogel actuators with larger actuation forces. Hydrogels actuated by pressured water (*6*) deliver a much higher actuation force (≈2 N) than existing osmotic-driven counterparts (≤0.01 N). Meanwhile, inspired by the superior leap ability of frogs, a supramolecular hydrogel (*7*) exhibits an actuation force of 0.3 N by storing and releasing elastic potential energy. However, strong actuation forces are needed to realize their real-world applications in soft robotics, where they are often required to perform laborious mechanical tasks.

Another bottleneck encountered by osmotic pressure–driven hydrogel actuators is their slow actuation speed owing to the diffusion-limited water transport. One general approach to increase the actuation rate is by the introduction of pores through freeze-thaw (*8*), three-dimensional (3D) printing (*9*), and phase transitions (*10*). Emerging actuation mechanisms that do not rely on water diffusion, such as electrostatic permittivity change (*11*) and light-induced bubble formation (*12*), can enable not only ultrafast responses (≤1 s) but also actuation in an open-air environment. Although their actuation speed is acceptable for most practical applications, more improvements are still needed to achieve high actuation force while retaining fast responses.

In nature, plant cells retain a high turgor pressure because of the confinement effect of cell walls on transported water. Inspired by this phenomenon, Na *et al.* created a confined swelling environment by wrapping a hydrogel with a permeable and stiff membrane, which resulted in an ultrahigh osmotic pressure ($\pi_{os}^{wrapped}$) (see the figure). Theoretical calculation revealed a negligible contribution of polymer elastic stress (σ_{el}), which prevented actuation. Both factors contributed to an ultrastrong actuation force (730 N), which is three orders of magnitude higher than that of existing hydrogel actuators. The turgor actuator could withstand a high compressive force of 917 N without fracture. This allowed it to break a rigid brick, which is impossible for current hydrogels.

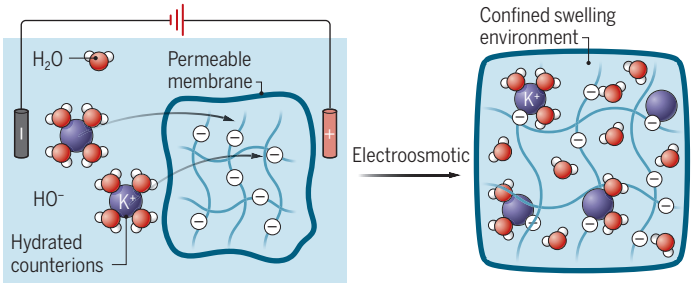
To further boost the actuation speed, an electric field was applied to drive hydrated counterions that accelerate the water migration to swell the network. This electrical-driven water transport had an actuation speed 19 times that of the corresponding osmotic rates and augmented the actuation forces. Turning on and off the field allowed for a reversible actuation over 20 cycles without any deterioration.

Na *et al.* open an exciting avenue for maximizing actuation force in hydrogels. Theoretical analyses provide guidelines for rationally designing and better understanding material performances. A turgor hydrogel that combines ultrahigh actuation force, high compressibility, and fast response will likley help to expedite the next generation of aquatic soft robotics capable of withstanding high underwater pressure.

Despite substantial progress, these materials are at an early stage. Future endeavours should be dedicated to realizing their excellent water-retention ability to function under nonaqueous conditions. The combination of surface modifications and innovative materials design could be a promising direction for the next generation of integrated smart hydrogels exhibiting fast, reversible, and high-powered actuation in multiple environments.

Turgor design for a high-power actuator

The hydrogels were designed to retain high osmotic pressure by wrapping with a stiff but permeable membrane to create a confined swelling environment. To improve actuation speed, an electric field is applied that drives hydrated counterions (left) into the hydrogel (right). This accelerates the hydrogel swelling.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

[1]Centre for Future Materials, University of Southern Queensland, Springfield Central, QLD 4300, Australia. [2]School of Agriculture and Environmental Science, University of Southern Queensland, Springfield Central, QLD 4300, Australia. Email: pingan.song@usq.edu.au

REFERENCES AND NOTES

1. I. Apsite, S. Salehi, L. Ionov, *Chem. Rev.* **122**, 1349 (2022).
2. X. Le, W. Lu, J. Zhang, T. Chen, *Adv. Sci. (Weinh.)* **6**, 1801584 (2019).
3. H. Na *et al.*, *Science* **376**, 301 (2022).
4. S. Y. Zheng *et al.*, *Adv. Funct. Mater.* **28**, 1803366 (2018).
5. M. Hua *et al.*, *ACS Appl. Mater. Interfaces* **13**, 12689 (2021).
6. H. Yuk *et al.*, *Nat. Commun.* **8**, 14230 (2017).
7. Y. Ma *et al.*, *Sci. Adv.* **6**, eabd2520 (2020).
8. Z. Jiang, B. Diggle, I. C. G. Shackleford, L. A. Connal, *Adv. Mater.* **31**, 1904956 (2019).
9. S. M. Chin *et al.*, *Nat. Commun.* **9**, 2395 (2018).
10. Z. Jiang *et al.*, *Chem. Mater.* **33**, 7818 (2021).
11. Y. S. Kim *et al.*, *Nat. Mater.* **14**, 1002 (2015).
12. M. Li, *et al.*, *Nat. Commun.* **11**, 3988 (2020).

ACKNOWLEDGMENTS

The authors acknowledge funding from the Australian Research Council (nos. FT190100188 and DP190102992).

10.1126/science.abo4603

SEISMOLOGY

Citizen science for studying earthquakes

Seismologist-citizen partnership helped understand the 2021 Haiti earthquake

Firefighters remove debris in search of survivors after the August 2021 earthquake in Haiti. First responders, such as the ones shown here, will benefit from the improved earthquake monitoring provided by the citizen-science Raspberry Shake network.

PHOTO: RICHARD PIERRIN/GETTY IMAGES

By **Christa von Hillebrandt-Andrade**[1] *and* **Elizabeth Vanacore**[2]

At 8:29 a.m. local time [12:29 universal time coordinated (UTC)] on 14 August 2021, the furthest thing from the mind of Haitians was another devastating earthquake. Many had thought that after the 12 January 2010 earthquake [moment magnitude (M_w) 7.0], they would have a respite from this hazard. In the end, the 2021 quake was even more powerful (M_w 7.2), releasing ~40% more energy than the 2010 earthquake (*1*). Tragically, the earthquake killed 2246 people, injured 12,763, left 329 missing, and affected at least 800,000 more people, 650,000 of whom required emergency humanitarian assistance. In addition, water, sanitation, and health facilities were all severely impaired (*2*). The impact was compounded because of the sociopolitical and economic challenges plaguing the country. On page 283 of this issue, Calais *et al.* (*3*) present a case study in the application of citizen science in real-time earthquake monitoring, response, and scientific inquiry.

Haiti is located on the western portion of the island of Hispaniola on the Caribbean plate that is bounded by the North American plate to the North. The Caribbean plate and the North American plate converge obliquely at a rate of ~2 cm/year, with the Caribbean tectonic plate moving east relative to the North American plate (*4*). The Puerto Rico Trench, which is the deepest part of the Atlantic Ocean and the Caribbean Sea, along with the North Hispaniola Fault and the Septentrional Fault to the north and the Enriquillo Fault to the south (*5*), accommodate the strain where the plates converge. These fault systems have consistently generated very large earthquakes and tsunamis based on modern seismic monitoring and historical accounts from the past ~500 years (*4*).

As with most earthquakes, the first questions surrounding the 2021 quake were about its strength, its epicenter, and the possibility of a tsunami. In most countries, a national seismic network or tsunami warning center could provide the answers. However, at the time of this earthquake, the Haitian Seismic Network was not operational, and only five seismometers in the entire country recorded the earthquake—specifically, three citizen-hosted sensors, one sensor at the US embassy, and one at a local high school.

The alarm system at the Pacific Tsunami Warning Center (PTWC) in Hawaii—the designated tsunami service provider for the Caribbean and adjacent regions—was triggered by seismic signals emanating from stations in Cuba and the Dominican Republic, which were about 200 and 240 km from the epicenter, respectively. Within 10 min of the earthquake, 5 min slower than the goal response time, the PTWC issued the first bulletin with the earthquake epicenter located 120 km west of Port-au-Prince, Haiti. The notification indicated no tsunami threat based on the initially estimated magnitude of 7.0 (*6*).

Nine minutes later, the US Geological Survey (USGS) issued a preliminary analysis of data, which included more regional and global stations, and determined the magnitude of the earthquake to be 7.2 (*1*). This larger magnitude then triggered the PTWC to issue a tsunami threat message for Haiti at 9:14 a.m. local time (*6*). By that time, many people had already self-evacuated, given the strong ground shaking. However, in response to the PTWC message, national authorities issued an official tsunami warning, prompting additional evacuations (*7*). The PTWC measured a tsunami of only 2 cm at Port-au-Prince at 10:00 a.m. local time. It then issued a final threat message at 10:19 a.m. local time (*6*), after which the warning was canceled (*7*).

Access to nearer field seismic and sea level data could have resulted in a more rapid analysis of the earthquake and an earlier warning of a potential tsunami threat by the PTWC. This was the case for the citizen-science network, which integrated data from its stations closer to the epicenter, as well as regional data, and published the earthquake's size and location within a few minutes (*8*). The parameters, both location and magnitude, that were calculated by the network were comparable to those of the organizations using regional data.

The citizen-science network included 15 plug-and-play low-cost sensors dubbed Raspberry Shakes (RSs), which are class C sensors according to the US Advanced National Seismic System (*8*, *9*). National and international seismologists established the RS network in Haiti in 2019, given the challenges with the national seismic system (*10*). The program supplies the sensors to private individuals, who in turn provide the electricity and internet. Along with the RS network, seismologists also developed a data-sharing platform named Ayiti-Séismes that integrates data from the RS network and other national and regional seismic data to automatically calculate and display the location and magnitude of local earthquakes (*10*).

Smaller earthquakes, called aftershocks, follow large earthquakes. The more sensory stations near an earthquake source, the more sensitive the overall sensory network. A more sensitive network would have a lower magnitude threshold and thus detect and report more earthquakes. Based on the data collected by the RS system, 1031 aftershocks were located within the first 3 weeks after the 2021 mainshock. By comparison,

[1]International Tsunami Information Center–Caribbean Office, Honolulu, HI, USA. [2]Puerto Rico Seismic Network, Department of Geology, University of Puerto Rico at Mayagüez, Mayagüez, Puerto Rico. Email: christa.vonh@noaa.gov

the USGS only reported 37 aftershocks in the same time period.

Aftershocks can also be forecasted, but this depends upon the timely and accurate detection of earthquakes; lowering the magnitude threshold of the data used in the forecast model leads to improved aftershock forecasting. The RS network in Haiti demonstrates that these class C sensors can improve earthquake data catalogs that serve as inputs for aftershock forecasting.

Another important question to answer is what fault or faults are responsible for the shaking. With the earthquake locations obtained using the RS data, Calais *et al.* determined the source of the activity to be an east-west fault zone ~80 km long with seismicity concentrated in two clusters. They also determined that the cluster to the east, which included the mainshock, was associated with vertical motion along the Enriquillo Fault. Meanwhile, the second cluster further to the west along the Ravine du Sud Fault was associated with lateral motion. The identification of the clusters was possible thanks to the lower magnitude detection threshold of the citizen-science network, thus demonstrating that low-cost sensors can also provide valuable scientific information.

The use and expansion of the low-cost class C sensors will not replace the need for national and regional seismic networks but do provide an avenue to expand network coverage in regions with logistical, economic, geographical, or other challenges that limit possible installation of class A and B sensors. The science-public partnership and the expanded use of RSs or similar instruments, such as in Haiti, also provide a possible avenue to expand earthquake-monitoring capabilities to underserved communities to foster disaster risk reduction. ■

REFERENCES AND NOTES

1. USGS, M 7.2 – Nippes, Haiti; https://earthquake.usgs.gov/earthquakes/eventpage/us6000f65h/executive.
2. Reliefweb, Haiti: Earthquake situation report no. 8 – Final (29 November 2021); https://reliefweb.int/report/haiti/haiti-earthquake-situation-report-no-8-final-29-november-2021.
3. E. Calais *et al.*, *Science* **376**, 283 (2022).
4. W. H. Bakun, C. H. Flores, U. S. ten Brink, *Bull. Seismol. Soc. Am.* **102**, 18 (2012).
5. S. Symithe, E. Calais, J. B. de Chabalier, R. Robertson, M. Higgins, *J. Geophys. Res. Solid Earth* **120**, 3748 (2015).
6. National Oceanic and Atmospheric Administration–National Weather Service, U.S. Tsunami Warning System; https://tsunami.gov [accessed 23 February 2022].
7. United Nations Educational, Scientific and Cultural Organization–Intergovernmental Oceanographic Commission (UNESCO/IOC) Caribbean Tsunami Information Center; https://www.ctic.ioc-unesco.org/ [accessed 18 February 2022].
8. E. Calais *et al.*, *Front. Earth Sci.* **8**, 122 (2020).
9. Working Group on Instrumentation, Siting, Installation, and Site Metadata of the Advanced National Seismic System Technical Integration Committee, "Instrumentation Guidelines for the Advanced National Seismic System" (Open-File Report 2008–1262, USGS, 2020); http://pubs.usgs.gov/of/2008/1262.
10. Ayiti-SEISMES; https://ayiti.unice.fr/ayiti-seismes/.

10.1126/science.abo5378

METABOLISM

Complex regulation of fatty liver disease

Hepatic lipogenesis is fine-tuned by mechanistic target of rapamycin (mTOR) signaling

By **Henry N. Ginsberg**[1] *and* **Arya Mani**[2]

Nonalcoholic fatty liver disease (NAFLD) is an umbrella term for hepatic abnormalities, including steatosis (fat accumulation, NAFL) and nonalcoholic steatohepatitis (NASH), which is NAFL plus hepatic injury, inflammation, and fibrosis (*1*). NAFLD has a prevalence of ~25% worldwide and results from the inability of the liver to maintain lipid homeostasis, leading to accumulation of triglyceride (TG), the major energy-storage molecule in mammals. Obesity, insulin resistance, and diabetes mellitus are drivers of NAFLD, so it is not surprising that mechanistic target of rapamycin (mTOR), which sits at the crossroads of nutrient signaling (*2*), plays a critical role in its etiology. It was also expected that the role of mTOR would be complex, but the extent of this complexity seems endless. On page 364 of this issue, Gosis *et al.* (*3*) present evidence that selective inhibition of a noncanonical arm of mTOR complex 1 (mTORC1) signaling inhibits hepatic de novo lipogenesis (DNL) and protects mice from NAFLD.

The normal liver contains between 15 and 75 g of hepatic TG (1 to 5% of an ~1500-g total liver weight). In NAFL, liver fat may increase to 20 to 30% of a 2000-g liver, or ~500 g of hepatic TG. Steatosis can lead to substantial hepatic pathology, including cirrhosis and hepatocellular carcinoma. Four major metabolic processes regulate hepatic TG amounts. The major driver of TG accumulation, which accounts for 65 to 70% of hepatic TG, is delivery of plasma albumin-bound fatty acids (FAs), which are derived mainly from adipose tissue (*4*). The second pathway for accumulation is DNL, which is the synthesis of TG from acetyl–coenzyme A derived mainly from metabolism of glucose in the mitochondria. DNL can account for 5 to 30% of hepatic TG (*4*, *5*). The two pathways responsible for "disposing" of hepatic TG and maintaining normal hepatic TG content are oxidation of FAs and secretion of TG in very-low-density lipoprotein (VLDL). All of these pathways, which are altered in individuals with insulin resistance, obesity, and diabetes mellitus, are regulated, at least in part, by mTOR.

mTOR was identified in the mid-1990s as a protein kinase that was the target of the immunosuppressive drug rapamycin, when in complex with 12-kDa FK506-binding protein (FKBP12) (*2*, *6*, *7*). Subsequently, the involvement of mTOR in many central cellular functions beyond immunosuppression was identified, as were two key regulatory components. mTOR exists in two distinct complexes: regulatory-associated protein of mTOR (RAPTOR) "defines" mTORC1, and rapamycin-insensitive companion of mTOR (RICTOR) defines mTORC2. There is a detailed understanding of the regulation of each mTORC by numerous proteins as well as the many downstream processes regulated by each, depending on signals from hormones, nutrients, and energy-producing pathways (*2*). The number of molecules involved, as well as the many autoregulatory feedback loops, suggests that there are more molecules and pathways left to be discovered, as in the case of NAFLD and DNL.

A link between insulin signaling and the sterol-regulatory element binding proteins (SREBPs), particularly SREBP-1c, was demonstrated in the late 1990s (*8*, *9*). Subsequent studies showed that insulin signaling, through its hepatic receptor, is required for the proteolytic processing and transport to the nucleus of SREBP-1c, where it transcriptionally activates several genes required for DNL. Further studies generated inconsistent and sometimes conflicting data regarding the regulation of DNL by mTORC1. For example, studies indicated that deletion of *Raptor*, which reduced mTORC1 activity, or deletion of tuberous sclerosis complex 1 (*Tsc1*) or *Tsc2*, which activated mTORC1, both resulted in reduced DNL and protection from hepatic steatosis in mice (*10*, *11*). Gosis *et al.* attempted to clarify these conflicting data. They identified a noncanonical pathway involving the protein folliculin (FLCN) that, when depleted in livers of mice, results in suppressed SREBP-1c activity and DNL, with protection against NAFLD.

[1]Department of Medicine, Irving Institute for Clinical and Translational Research, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA. [2]Department of Medicine, Yale Cardiovascular Research Center, Yale University School of Medicine, New Haven, CT, USA. Email: hng1@cumc.columbia.edu

The mechanism they uncovered involves activation of the canonical mTORC1-S6 kinase (S6K) pathway, leading to "feedback" down-regulation of FLCN and suppression of mTORC1-mediated phosphorylation of the transcription factor TFE3 by a noncanonical pathway. This results in translocation of unphosphorylated TFE3 into the nucleus, where it increases expression of insulin-induced gene 2 (*Insig2*), ultimately leading to inhibition of the proteolytic processing of SREBP-1c in the endoplasmic reticulum and Golgi apparatus that is necessary for its transcriptional activity (see the figure). This finding is consistent with the demonstration that activation of mTORC1 by disruption of TSC1 and TSC2 is insufficient for stimulation of DNL and requires suppression of *Insig2* (*11*).

Although Gosis *et al.* used mice in which *Flcn* was ablated for most of their work, they raised the possibility that targeting FLCN could be an effective therapy to prevent or diminish NAFLD. Targeting FLCN could avoid inhibiting targets of the canonical mTORC1 pathway, including p70S6K, which can both stimulate and inhibit DNL. It is unlikely, based on the complexity of the metabolic pathways regulated by mTORC1, that inhibiting FLCN, and thereby activating TFE3, alone would completely inhibit DNL.

There are additional insights into the complexity of hepatic mTOR signaling. The activity of mTORC1 was found to be important for the maintenance of VLDL secretion by increasing the expression of CTP:phosphocholine cytidyltransferase α (CCTα), a key enzyme in the generation of phosphatidylcholine, which is required for VLDL assembly and secretion (*12*). Another study focused on mTORC2 and the dual-specificity tyrosine phosphorylation-regulated kinase (DYRK1B), which has been linked to metabolic syndrome in several kindreds (*13*). Increasing DYRK1B expression in the livers of mice that were fed high-fat, high-sucrose diets increased hepatic DNL, FA uptake, and TG secretion, concomitant with development of hyperlipidemia and NASH (*14*). Disruption of mTORC2 reversed these abnormalities.

Knowledge about the role of each mTORC and their components in the regulation of DNL and the development of NAFLD remains far from complete. Additional studies, hopefully not never-ending, are needed for the development of therapies that can target components of mTOR and prevent the development of NAFLD without inhibiting the many critical roles of this master regulator of cell metabolism. ■

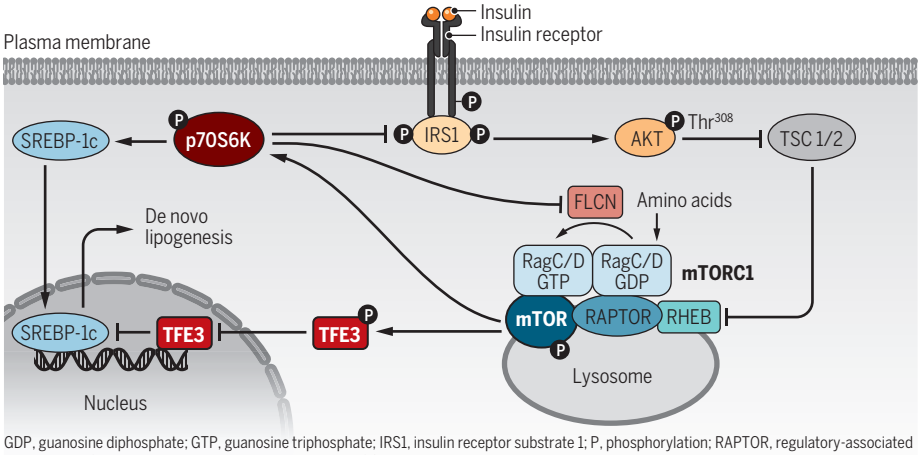
REFERENCES AND NOTES

1. R. Loomba, S. L. Friedman, G. I. Shulman, *Cell* **184**, 2537 (2021).
2. G. Y. Liu, D. M. Sabatini, *Nat. Rev. Mol. Cell Biol.* **21**, 183 (2020).
3. B. S. Gosis *et al.*, *Science* **376**, eabf8271 (2022).
4. K. L. Donnelly *et al.*, *J. Clin. Invest.* **115**, 1343 (2005).
5. G. I. Smith *et al.*, *J. Clin. Invest.* **130**, 1453 (2020).
6. E. J. Brown *et al.*, *Nature* **369**, 756 (1994).
7. C. J. Sabers *et al.*, *J. Biol. Chem.* **270**, 815 (1995).
8. J. D. Horton, Y. Bashmakov, I. Shimomura, H. Shimano, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5987 (1998).
9. M. Foretz *et al.*, *Mol. Cell. Biol.* **19**, 3760 (1999).
10. T. R. Peterson *et al.*, *Cell* **146**, 408 (2011).
11. J. L. Yecies *et al.*, *Cell Metab.* **14**, 21 (2011).
12. W. J. Quinn 3rd *et al.*, *J. Clin. Invest.* **127**, 4207 (2017).
13. A. R. Keramati *et al.*, *N. Engl. J. Med.* **370**, 1909 (2014).
14. N. Bhat *et al.*, *J. Clin. Invest.* **132**, e153724 (2022).

10.1126/science.abp8276

Complex regulation of de novo lipogenesis

Insulin activates the insulin receptor on hepatocytes, which activates the mechanistic target of rapamycin (mTOR) complex 1 (mTORC1) and folliculin (FLCN). FLCN inhibits transcription factor E3 (TFE3) nuclear localization and its suppression of sterol-regulatory element binding protein 1c (SREBP-1c) and de novo lipogenesis through a noncanonical pathway. Disruption of FLCN allows TFE3 to enter the nucleus and inhibit SREBP-1c without affecting p70 S6 kinase (p70S6K) activity, including its negative feedback of insulin signaling.



GDP, guanosine diphosphate; GTP, guanosine triphosphate; IRS1, insulin receptor substrate 1; P, phosphorylation; RAPTOR, regulatory-associated protein of mTOR; RHEB, Ras homolog enriched in brain; TSC, tuberous sclerosis complex.

PHYSIOLOGY

How the gut talks to the brain

Peptidoglycans from gut microbiota modulate appetite through hypothalamic circuits

By **Antoine Adamantidis**[1,2]

The mammalian gastrointestinal tract hosts a community of diverse microorganisms, including bacteria, archea, fungi, and viruses. Bacterial products, such as metabolites and cell wall fragments, are implicated in host metabolic functions. In addition, the gut microbiota influences the immune and central nervous systems, and it has emerged as a key regulator of brain development and the modulation of behaviors, including stress and anxiety, often in a sex-specific manner (*1*). Disruption of gut microbiota–brain interactions contribute to the pathogenesis of neurodevelopmental and psychiatric disorders in animal models (*2*). On page 263 of this issue, Gabanyi *et al.* (*3*) show that bacterial peptidoglycans, a by-product of bacterial cell wall degradation during cell division and cell death, directly inhibit the activity of feeding-promoting neurons in the hypothalamus and ultimately decrease appetite and body temperature, mostly in female mice. This finding may open new approaches for the treatment of metabolic disorders, including obesity.

The gut microbiota is a source of diverse bacterial peptidoglycan fragments. Peptidoglycan is an essential component of the bacterial cell wall that is absent from eukaryotic cells. It is a rigid and insoluble polymer composed of glycan strands cross-linked by peptides. Peptidoglycans have also been an important target for antibacterial drug discovery and development, and it is now a potential therapeutic target for metabolic and mental health disorders (*4*).

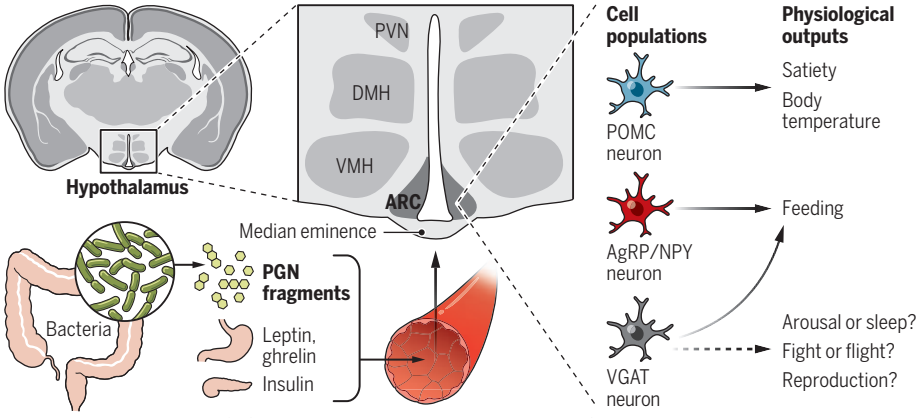
Direct and indirect pathways support gut–brain communications, including neural signals (e.g., the vagus nerve relays peripheral

[1]Zentrum für Experimentelle Neurologie, Department of Neurology, Inselspital University Hospital Bern, Bern, Switzerland. [2]Department of Biomedical Research, University of Bern, Bern, Switzerland.
Email: antoine.adamantidis@dbmr.unibe.ch

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

The gut microbiota–hypothalamus connection

Peptidoglycan (PGN) fragments from bacterial cell walls circulate and reach the brain, where they decrease the activity of feeding-promoting agouti-related peptide/neuropeptide-Y (AgRP/NPY) neurons in the arcuate nucleus (ARC) and thus food intake. This adds to the multiple peripheral signals that are detected by hypothalamus cells to control appetite and energy homeostasis, as well as other physiological outputs.



DMH, dorsomedial hypothalamus; POMC, pro-opiomelanocortin; PVN, paraventricular nuclei; VGAT, vesicular γ-aminobutyric acid transporter; VMH, ventromedial hypothalamus.

signals, including from the intestines, to the central nervous system), immune factors (e.g., circulating cytokines and chemokines released within the gut), and signaling hormones or metabolites. Microbial products, such as metabolites, fatty acids (acetate, propionate, and butyrate), and even neurotransmitters (dopamine and serotonin), influence brain cells, including neurons and microglia, in multiple brain structures. Hence, it is not surprising that the alteration of these communication pathways is associated with a range of neurodevelopmental and psychiatric disorders (*1*, *4*).

Although the signaling pathways involved in gut-brain communication remain to be fully elucidated, peptidoglycans have been detected in the bone marrow, blood, and cerebrospinal fluid (CSF) of healthy humans, rodents, and nonhuman primates, where specific receptors or other proteins can detect their presence (*1*, *3*). Indeed, peptidoglycans bind to pattern-recognition receptors (PRRs), cytosolic nucleotide-binding oligomerization domain-like receptors (NOD1 and NOD2), and membrane-bound Toll-like receptors (TLRs) and peptidoglycan recognition proteins (PGRPs). Gabanyi *et al.* show that NOD2 is expressed by neurons in the mouse brain, including the hypothalamus, in a region- and sex-dependent manner. They further show that intestinal bacteria–derived peptidoglycans called muropeptides (MDPs) bind to NOD2 and decrease the activity of inhibitory neurons from the arcuate nucleus (ARC) and ultimately decrease appetite in old female mice.

The ARC lies at the floor of the hypothalamus and is adjacent to the median eminence, a structure that is permeable to blood-borne metabolites, hormonal signals, and presumably bacterial products (*5*). The ARC contains agouti-related peptide/neuropeptide-Y (AgRP/NPY)–expressing neurons that signal hunger and promote feeding, which Gabanyi *et al.* also found to express the vesicular γ-aminobutyric acid transporter (VGAT) and NOD2 and are inhibited upon exposure to MDPs. Deletion of *Nod2* in VGAT$^+$ neurons in the ARC or the dorsomedial hypothalamus (DMH) resulted in reduced weight, feeding behavior, and nest-building; impaired temperature regulation; and decreased life span in female mice (see the figure).

In contrast to hunger AgRP/NPY neurons, pro-opiomelanocortin (POMC)–expressing neurons signal satiety, and their activation stops feeding. These two neuronal populations act as "first-order" sensors of energy status and integrate peripheral signals, including the hormones leptin, ghrelin, and insulin, to regulate feeding and energy expenditure. Thus, besides MDP inhibition of feeding-promoting cells (AgRP/NPY VGAT$^+$) reported by Gabanyi *et al.*, it remains unclear whether other microbiota-derived molecules modulate POMC neurons or other hypothalamic cells, including tanycytes, a type of glial cell that modulates food intake and energy homeostasis. In this context, second-order neural circuits of the lateral hypothalamus that act as metabolic sensors and control food intake or energy homeostasis, some of which include feeding-promoting VGAT$^+$ neurons (*6*), may represent other possible peptidoglycan targets that warrant further investigation.

In addition to regulating energy homeostasis, the hypothalamus is critical for fight-or-flight responses (including stress and anxiety), reproduction, sleep-wake states, and goal-directed behaviors toward natural (food, sex) and artificial (drug) rewards (*7*). It is a federation of nuclei that encompass multiple cell populations with complex neurochemical profiles and electrophysiological fingerprints that form an intricate local and extensive network of excitatory and inhibitory cells, each of which has a specific role in homeostatic functions. Notably, the hypothalamus modulates other vital functions, including the quantity and the quality of sleep (*8*), both of which are modulated by microbial products (*9*). Peptidoglycan fragments are detected in the CSF and brain of sleep-deprived animals and urine of sleep-deprived humans. Consistently, administration of peptidoglycan fragments increased the duration of non–rapid eye movement (NREM) sleep in rodents and nonhuman primates (*10*) and enhanced the amplitude of slow waves—a marker of sleep pressure and sleep quality—during episodes of NREM sleep. Similarly, such modulation of hypothalamic VGAT$^+$ neurons may be involved in arousal control because these neurons have also been implicated in the control of wakefulness (*11*).

The microbiota provides potential biological markers of microbiota-brain interactions and candidates for the development of therapeutic strategies for the treatment of neurodevelopmental, psychiatric, and metabolic disorders. A major challenge in identifying these mechanisms is the multiple brain targets and the diversity of hypothalamic cellular pathways. Thus, a better understanding of the cellular cross-talk between appetite and body temperature, as well as other hypothalamic circuits that share functions—including those that regulate sleep and body temperature (*12*), sleep and appetite (*13*), reproductive and aggressive behaviors (*14*), or arousal and reward processing (*15*)—will be essential to reveal the mechanisms and therapeutic relevance of potential treatments. ■

REFERENCES AND NOTES

1. A. Gonzalez-Santana, R. Diaz Heijtz, *Trends Mol. Med.* **26**, 729 (2020).
2. J. Nagpal, J. F. Cryan, *Neuron* **109**, 3930 (2021).
3. I. Gabanyi *et al.*, *Science* **376**, 263 (2022).
4. K. Berding, J. F. Cryan, *Curr. Opin. Psychiatry* **35**, 3 (2022).
5. A. Jais, J. C. Brüning, *Endocr. Rev.* **43**, 314 (2022).
6. J. H. Jennings *et al.*, *Cell* **160**, 516 (2015).
7. G. D. Stuber, R. A. Wise, *Nat. Neurosci.* **19**, 198 (2016).
8. A. Adamantidis, L. de Lecea, *Trends Endocrinol. Metab.* **19**, 362 (2008).
9. M. Sgro *et al.*, *Sleep* **45**, zsab268 (2022).
10. J. M. Krueger, J. R. Pappenheimer, M. L. Karnovsky, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6102 (1982).
11. C. G. Herrera *et al.*, *Nat. Neurosci.* **19**, 290 (2016).
12. T. M. Takahashi *et al.*, *Nature* **583**, 109 (2020).
13. L. T. Oesch *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 19590 (2020).
14. H. Lee *et al.*, *Nature* **509**, 627 (2014).
15. T. Sakurai, *Nat. Rev. Neurosci.* **15**, 719 (2014).

ACKNOWLEDGMENTS

A.A. is supported by the Swiss National Science Foundation, the European Research Council, the Inselspital, and the University of Bern.

10.1126/scienceBecky

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

POLICY FORUM

GENETICS AND SOCIETY

Getting genetic ancestry right for science and society

We must embrace a multidimensional, continuous view of ancestry and move away from continental ancestry categories

By **Anna C. F. Lewis, Santiago J. Molina, Paul S. Appelbaum, Bege Dauda, Anna Di Rienzo, Agustin Fuentes, Stephanie M. Fullerton, Nanibaa' A. Garrison, Nayanika Ghosh, Evelynn M. Hammonds, David S. Jones, Eimear E. Kenny, Peter Kraft, Sandra S.-J. Lee, Madelyn Mauro, John Novembre, Aaron Panofsky, Mashaal Sohail, Benjamin M. Neale, Danielle S. Allen**

Glaring health disparities have reinvigorated debate about the relevance of race to health, including how race should and should not be used as a variable in research and biomedicine (*1*). After a long history of race being treated as a biological variable, there is now broad agreement that racial classifications are a product of historically contingent social, economic, and political processes. Many institutions have thus been reexamining their use of race and racism and stating intentions about how race should be used going forward. One common proposal is to use genetic concepts—in particular, genetic ancestry and population categories—as a replacement for race (*2*). However, the use of ancestry categories has technical limitations, fails to adequately capture human genetic diversity and demographic history, and risks retaining one of the most problematic aspects of race—an essentialist link to biology—by allowing genetic ancestry categories to stand in its place.

The process of racialization entails a dynamic cognitive process of identification based on phenotype that is often highly context dependent. Although research has found genetic variation correlated with phenotypes that have been historically used to assign race categories, such as skin pigmentation or hair texture, it is the case that such genetic correlates are not distributed in a manner that correspond to racially defined groups. Race is a sociopolitical construct rather than a biological one. For example, in the United States, immigrants from southern and eastern Europe only began to be classified as "white" on the census in the 20th century (*3*); the American Indian/Alaska Native census category reflects colonizing histories and federal policies (*4*). As such, social scientists and others have argued that the strongest case for using race is limited to tracking the impact of racism on health outcomes, rather than as a proxy for anything biological (*5*).

Genetic ancestry, one of the main proposed alternatives to using race, is of relevance to statistical and population geneticists, epidemiologists, public health practitioners, physicians, and patients. In particular, genetic ancestry has renewed relevance for the clinical application of genetic technology because the accuracy of genetic risk scores varies across ancestries (*6*). Genetic ancestry and population categories are also relevant to the general public, as demonstrated by the tens of millions of individuals who have paid for ancestry reports from consumer companies. Across these different domains, a dominant description of genetic ancestry is associated with continents as meaningful groupings. Within genetics research, continental ancestry categories have become the most common type of group label (*7*). Similarly, consumer genetics products give customers a report with data based on a percentage of these continental groups from which an individual can trace their "ancestry."

Systems of racial classification have historically regarded continents as meaningful group boundaries; thus, it is not surprising that racial categories and continental ancestry categories are often confounded. Whenever continental ancestry categories are used, the risk is high that a misconception of race as a biological attribute will reenter through the back door (*8*). Insufficiently nuanced thinking about continental categories, genetic ancestry, and racial groups can lead to the conflation of the three.

A FLATTENED NOTION OF ANCESTRY

Our genetic ancestry is defined by the stretches of the genome that we inherit from our ancestors (*9*). Geneticists have a concept for this known as the ancestral recombination graph (ARG). Put simply, an individual's genetic ancestry is the subset of paths through the human family tree by which they have inherited DNA from specific ancestors. Most often, geneticists study the ARG of multiple individuals at the same time.

Crucially, this definition makes clear that there are two things that are not necessary to the definition of genetic ancestry. The first is any categorization by populations or groups. And the second is any contextualization of the individuals apart from their genealogical connections—for example, by labeling these individuals with geographical or cultural information. Yet current practices around ancestry estimation and reporting almost always impose categories and, when they do so, very often default to just one way to contextualize individuals: by continent of origin. Both practices limit the accuracy and reliability of claims being made by researchers about human genetic difference.

There are many statistical methodologies across subfields of genetics and genomics whose outputs are framed as "genetic ancestry," most of which do not attempt to approximate the ARG and several of which only capture genetic similarity (*9*). The majority of these methods involve placing individuals into categories or modeling them as mixtures of discrete categories. For some methods, the categories are predefined and prelabeled. For others, the categories emerge from the analysis. In these cases, not only are the resulting categories very sensitive to which individuals are included in the analysis, they may not even represent shared ancestries (*10*). In other cases, categories and their labels are imposed in downstream analysis.

The concern about use of categories goes beyond these technical limitations. Imposing categories on genetic ancestry fails to adequately capture human genetic diversity and what we know of human demographic history. A standard way to visualize patterns of genetic similarity is by plotting results of principal components analysis of genetic variation data, a technique that reduces the dimensionality of that data. Most genetic analyses use data from reference populations to contextualize a study's data. The most commonly used reference data were created by sampling individuals from a few dozen places spread across the globe. If individuals from these populations are graphed in this manner, distinct clusters roughly representing continental categories are visible (see the figure). A prominent early result was that genetic ancestry was strongly concordant with continental origins when ascertaining for individuals

Author affiliations are available in the supplementary materials. Email: annalewis@g.harvard.edu

whose four grandparents were from the recruitment sites (*11*).

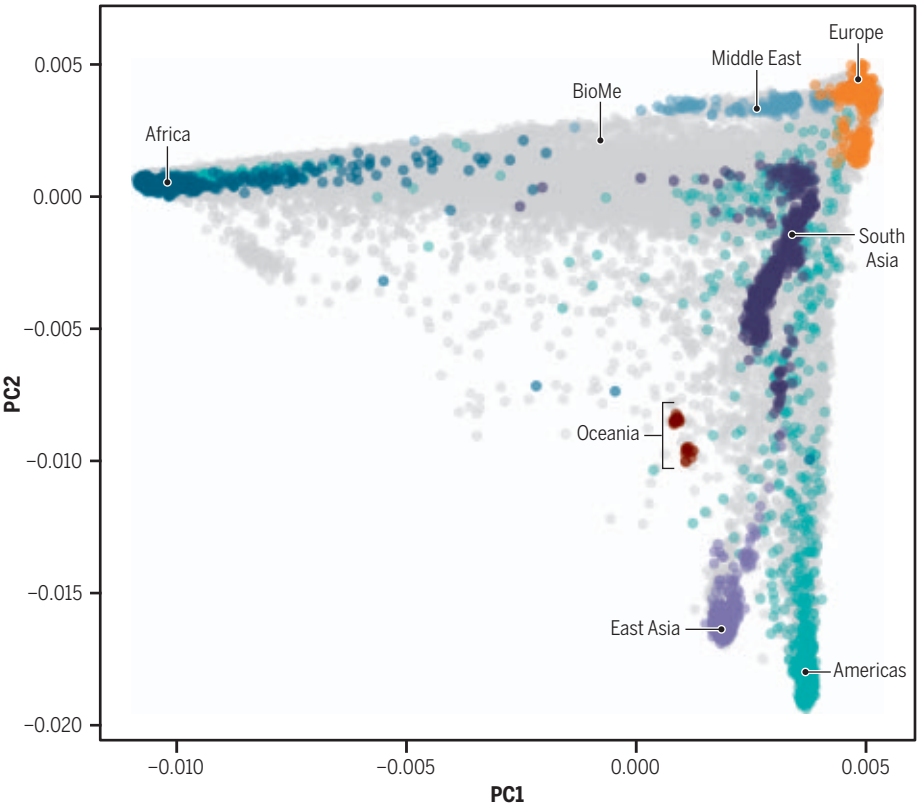
But newly assembled datasets show that if people are sampled differently, such as individuals living in New York City, it becomes clear how impoverished this view of a structure of distinct clusters is (see the figure) (*12*). The clearly separated clusters of reference population individuals, corresponding to different continental groups, merge into a background of continuous genetic variation. This is consistent with what we know of human demographic history, in which mass migration and constant mixing across groups have been the norm. The impact of these histories leads to different structures of genetic variation in different parts of the world. Such studies illustrate just how inappropriate use of discrete continental categories can be, particularly when information framed as genetic ancestry can potentially influence medical care.

The use of the terms admixture and "admixed individuals"—defined as those who have recent ancestry from more than one population, and typically continental ancestry populations—reinforces notions of discrete categories within humanity. This use does not escape the notion of continental ancestry categories but rather compounds the errors of using such categories because these individuals are typically conceptualized as a mixture of otherwise "pure" continental ancestry populations.

Our conceptualization of ancestry must be general enough to describe every human; the only way to do this is to use concepts and tools that acknowledge that ancestry is continuous. Categories have their legitimate uses—for example, in reporting the differences in predictive power of genetic risk scores (even in this case, differences in performance are due to many factors, and focusing on only one factor such as ancestry can lead to essentializing differences between groups) (*6*). But the default appeal to any one set of categories risks essentializing those groups, making it more likely that differences between these abstract groups are treated as though they were concrete.

The continuous, category-free, nature of genetic variation

Colored dots (n = 4149) are reference panel individuals from 87 populations representing ancestry from seven continental or subcontinental regions projected onto the first two principal components (PC1 and PC2) of genetic similarity. Gray dots (n = 31,705) are participants from BioMe, a diverse biobank based in New York City. Clearly delineated continental ancestry categories (the islands of color) are shown to be a by-product of sampling strategy. They are not reflective of the diversity in this real-world dataset, which is made evident by the continuous sea of gray.



GRAPHIC: K. FRANKLIN/*SCIENCE* BASED ON (*12*)

In addition to not requiring the use of categories, the definition of genetic ancestry is silent on any aspect of the context of an individual's ancestors. Although the ancestral recombination graph does have structure, it does not by itself indicate anything about an individual's geographical location or their culture. Researchers face choices in whether and how to provide this context. Crucially, we can give multiple contexts depending on the time horizon considered because we each have ancestors from every generation in our species' past. Advances in ancient DNA and in population genetics are providing us with more and more information about population structure at different points in our histories. A contemporary human genome can hence increasingly give us visibility into the chronologically layered ancestral record for that person.

Yet this historical notion of genetic ancestry is flattened when just one set of categories is used. In the case of continental ancestry categories, their use reflects the assumption that at some specific point in time, humans were mostly divided into homogeneous groups by the natural geographical barriers between continents. This is a gross oversimplification of human history. It also obscures other time slices when different categories would be relevant—for example, ~50,000 years ago, *Homo Sapiens* and Neanderthal categories; or ~5000 years ago, "Steppe-related," "European" hunter-gatherer, and "Near Eastern" farmer categories in Europe (*13*); or ~500 years ago, when waves of migration and the slave trade were forging new patterns of human genetic diversity in the Americas.

A MORE COMPLEX NOTION OF ANCESTRY

What are the implications for researchers who want to invoke genetic ancestry? They should first ask whether they need to impose categories at all to answer their research question. There are many situations in which categorization has been thought essential but has subsequently been shown to be avoidable, such as in correcting for population stratification in genome-wide association studies (*14*). In cases in which genetic ancestry categories can be avoided, they should be avoided. If researchers are able to justify a scientific need to impose categories, they should next think about whether they have to provide labels (be it geographic, ethnic, linguistic, or other) to the groupings they impose. If they do need to provide labels, they should give the scientific justification for that choice and show that they have considered potential disadvantages of imposing these labels.

Additionally, researchers should use multiple types of categories, reflecting that

genetic ancestry is a historical concept: We all have multiple ancestries depending on the time horizon considered. No individual has a single "ancestry"; the plural should always be used. Different geographical resolutions—for example, "Yoruban" versus "West African"—can serve as proxies for different time slices. Ancestry categories from different time points may be of medical relevance. The incorporation of ancient DNA information can also allow for probing different time slices, although the promise of this approach will depend on how much ancient DNA can actually be recovered and analyzed. The use of continental ancestry categories as a proxy for one of the time slices considered must be particularly carefully justified because of the conflation of continental ancestry categories with racial groupings. Additionally, future work should find better ways to conceptualize the genetic ancestry of individuals whose recent ancestors come from distant parts of the ARG.

For some diseases that have a different prevalence in different populations, genetic risk factors may indeed be at play, a result of differences in the chance arrival of new mutations, demographic history, and historical environmental exposures. But although it is possible that genetics is playing a causal role in such cases, genetic ancestry may also be serving as a proxy for differences in environmental effects, including the effects of discrimination. Whenever researchers invoke any categories in understanding health outcomes, they need to make careful efforts to jointly model genetic and environmental effects and acknowledge that a failure to explain differences could be due to unmodeled factors.

Science is reductive, and a model that uses simple continental categories has been useful in starting the process of understanding human genetic diversity. But all models have their legitimate domains of application and limits, and a much more complex set of models should now be the norm across a wide variety of use cases. This is particularly important because although human genetics falls under the biological sciences, it is in fact a science at the intersection of several disciplines, including anthropology, demography, epidemiology, history, and sociology. Even if the limitations of models used are well understood by statistical and population geneticists, others may take the models to be descriptive of realities rather than recognizing that they merely formalize approximations and estimates, using reductive categories to do so. Hence, one of the risks of using these categories is that others may interpret them as true natural kinds, which is inaccurate. Instead, they are heuristics permitting the approximation or answering of very narrow sorts of questions. Because of the association of continental ancestry categories with racial groupings, this is particularly important for continental categories.

An individual researcher's use of continental ancestry categories is not in and of itself racist, but the cumulative impact of this practice has led to and sustains racism. Typological thinking about human difference has had damaging social consequences. Continued reliance on continental ancestry categories contributes to failures of inference, miscommunication between fields, and reported findings that are rooted in reductive and limited ways of understanding human difference. These are likely to exacerbate medical stereotypes about individuals and groups, contribute to health disparities rather than addressing them, and reify (mis)understandings of race as biological. Moreover, this problem is not limited to continental ancestry categories; national categories can and have been reified as biological for political goals (*15*).

The solution will require addressing the issues with how ancestry is conceptualized and used across the entire biomedical research ecosystem. This will involve the development, operationalization, and widespread use of a more complex notion of ancestry—one that disambiguates what is meant by genetic ancestry from related concepts, wherever possible does not treat ancestry as a categorical variable, and treats ancestry as reflecting a historical process, meaning that any study should use many different types of categories.

To aid this transition, a solid empirical understanding of how and why different fields use and operationalize the concept of ancestry is needed. To ensure that this more complex notion of ancestry is then used in practice will require systems-level change. New computational tools and data structures will be required—for example, a wider variety of proxies for genetic ancestry that do not impose categories, as well as easily accessible software tools to enable use of ancestry categories representing multiple time horizons. Further development and adoption of methodologies that directly estimate the ARG should be encouraged. Educational materials will need to be developed for scientists and physicians. Scientists of all stripes who engage in research that uses biological categories for humans should not work in isolation but as part of interdisciplinary teams, ideally including engagement with affected communities. In support of these efforts, journal editors should set standards, professional societies should publish best practices, and funders should carefully consider which research agendas they will support. It is paramount, as these organizations rightly critique the use of race as a biological variable, that use of continental ancestry categories does not become the new default. The US National Academies of Sciences, Engineering, and Medicine recently formed an ad hoc committee, "Use of Race, Ethnicity, and Ancestry as Population Descriptors in Genomics Research"; we are hopeful that this represents an opportunity for consideration and consolidation of the points raised here.

Adoption of a more complex notion of ancestry should in turn continue to inform the research agenda in population and statistical genetics and in ancient DNA research. It is in these fields, the home turf of the concept of genetic ancestry, that change in practice may have the largest overall impact. These changes are a prerequisite to any research that looks for connections between genetics and health disparities. More generally, with a more complex notion of ancestry that reflects continuous variation and historical depth, we can start to pave the way for a science that reflects the complex histories of human groups, including the power dynamics among them. ■

REFERENCES AND NOTES

1. D. A. Vyas, L. G. Eisenstein, D. S. Jones, *N. Engl. J. Med.* **383**, 874 (2020).
2. A. Oni-Orisan, Y. Mavura, Y. Banda, T. A. Thornton, R. Sebro, *N. Engl. J. Med.* **384**, 1163 (2021).
3. D. R. Roediger, *Working Toward Whiteness: How America's Immigrants Became White: The Strange Journey from Ellis Island to the Suburbs* (Basic Books, Text is Free of Markings ed., 2005).
4. E. A. Haozous, C. J. Strickland, J. F. Palacios, T. G. A. Solomon, *J. Environ. Public Health* **2014**, e321604 (2014).
5. J. H. Fujimura, T. Duster, R. Rajagopalan, *Soc. Stud. Sci.* **38**, 643 (2008).
6. A. R. Martin *et al.*, *Nat. Genet.* **51**, 584 (2019).
7. A. Panofsky, C. Bliss, *Am. Sociol. Rev.* **82**, 59 (2017).
8. T. Duster, *Backdoor to Eugenics* (Routledge, ed. 2, 2003).
9. I. Mathieson, A. Scally, *PLOS Genet.* **16**, e1008624 (2020).
10. D. J. Lawson, L. van Dorp, D. Falush, *Nat. Commun.* **9**, 3258 (2018).
11. N. A. Rosenberg *et al.*, *Science* **298**, 2381 (2002).
12. G. M. Belbin *et al.*, *Cell* **184**, 2068 (2021).
13. I. Olalde *et al.*, *Nature* **555**, 190 (2018).
14. G. L. Wojcik *et al.*, *Nature* **570**, 514 (2019).
15. W.-C. Sung, in *Asian Biotech: Ethics and Communities of Fate*, A. Ong, N. Chen, Eds. (Duke Univ. Press, 2010), pp. 263–288.

ACKNOWLEDGMENTS

This work was supported by National Institute of Mental Health administrative supplements 5000747-5500001474 to 3R37MH107649-06S1. B.M.N. and D.S.A. contributed equally to this work. A.C.F.L. owns stock in Fabric Genomics. E.E.K. has received personal fees from Regeneron Pharmaceuticals, 23&Me, and Illumina and serves on the advisory boards for Encompass Biosciences and Galateo Bio. B.M.N. is a member of the scientific advisory board at Deep Genomics and RBNC Therapeutics, a member of the scientific advisory committee at Milken, and a consultant for Camp4 Therapeutics and Merck.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm7530

10.1126/science.abm7530

BOOKS *et al.*

Drought and deforestation have transformed a once-fertile region of Madagascar into a dust bowl.

PHOTO: REUTERS/ALKIS KONSTANTINIDIS

CLIMATE POLICY

Confronting climate injustice

Social, racial, and economic disparities are crucial considerations in climate policies

By **Miriam Aczel**

While most people understand that we face a looming climate disaster characterized by severe droughts, melting glaciers, and increasingly common wildfires and superstorms, certain technical details, policy considerations, and related justice and equity issues remain murky. Two timely new books aim to fill these gaps in knowledge.

In *Is Science Enough? Forty Critical Questions About Climate Justice*, historian Aviva Chomsky breaks down the key concepts, terminology, and often-contentious debates that surround climate change so that audiences ranging from students to activists can easily understand them. As the title implies, Chomsky argues that scientific interventions are not sufficient to combat global warming. Our current economic paradigm, she argues, relies on "extracting and consuming the earth's resources in ever-increasing quantities, and turning them into waste," and such a system is incompatible with a healthy planet.

Moreover, she posits, the approach to development that has intensified since the Industrial Revolution is fundamentally unfair. Groups that have historically maintained control of fossil fuels have achieved greater growth and prosperity than others, often at the disadvantaged groups' expense. Many from the latter groups argue that they are due a piece of the development pie, but Earth's resources are inherently limited.

Chomsky's highly readable and accessible primer is divided into five sections that answer various questions on climate change, including how different sectors contribute to it; how renewable and zero-emission energy sources are defined; how the US subsidizes the fossil fuel industry; the relationship between social, racial, and economic justice and climate change; and how economic growth affects the environment. Although the book evaluates the efficacy of behavioral changes one might make to help combat climate change, such as buying electric cars, giving up flying, or changing to a meat-free diet, Chomsky is careful to emphasize that the problem is society-wide, meaning that effective interventions must necessarily transcend individual action.

She concludes that the climate problem reflects deeper issues of social injustice and inequity that cannot be fixed by "tweaking incentives and adding technologies." Major reorganization of the global economy and society will be necessary.

The reviewer is at the California Institute for Energy and Environment, University of California, Berkeley, CA 94720, USA. Email: aczel@berkeley.edu

In *What Climate Justice Means and Why We Should Care*, moral philosopher Elizabeth Cripps argues that we all share a responsibility to combat the effects of a changing climate that is disproportionately affecting those who have done the least to cause it. She presents clear and compelling evidence of the burden borne by disadvantaged populations, maintaining that climate change is, above all, "about privilege."

Ten countries, Cripps notes, are responsible for 60% of greenhouse gas emissions—a major driver of climate change—and while the impacts of climate change are global and include severe winter storms in Texas, wildfires in Australia, and floods in Europe, the Global South has suffered the most devastating consequences. Between 2008 and 2016, she writes, roughly 22 million people were displaced in the Global South each year. The consequences of such displacements include child marriages, loss of schooling and employment opportunities, food insecurity, and more.

Cripps concludes by arguing that we must all take action, as more than just the planet is in peril—we also need to ensure that we are "not killing our fellow human beings." Those most responsible for the damage have the greatest duty to act, however, as the suffering being wrought in marginalized communities is due to legacies of "colonialism, slavery, oppression, and systematic disregard for basic human rights." The book ends on a positive note, encouraging readers to remember that while we may at times feel powerless, power can be found in joining with others as a community of responsible global citizens.

We are facing a global existential threat that is exacerbated by the accelerating impacts of a changing climate and environmental destruction. How we respond will determine the future of life on Earth. As both Chomsky and Cripps show, addressing climate change is not just about devising technical or scientific solutions, it also requires acknowledging and addressing social, racial, and economic injustices that have played a role in the crisis. Both books, however, end on a cautiously positive note: If we put climate justice front and center and rethink how we view growth and the world economy, we can reach equitable and inclusive solutions to our changing climate. ■

What Climate Justice Means and Why We Should Care
Elizabeth Cripps
Bloomsbury Continuum, 2022. 224 pp.

Is Science Enough?
Aviva Chomsky
Beacon Press, 2022. 240 pp.

10.1126/science.abo3385

HEALTH AND MEDICINE

The complexity of chronic pain

A physician confronts an elusive physical phenomenon

By **Marcia L. Meldrum**

In *The Song of Our Scars: The Untold Story of Pain*, Haider Warraich confronts the enigma of chronic pain both as a physician and a patient. Building on his own experience with chronic pain—the result of an injury sustained during a weight-lifting accident while Warraich was in medical school—and extensive research and interviews with patients and scientists, he explores the neurophysiology of pain and explains why the nervous system mechanisms underlying acute pain fail to explain the persistence and intensity of chronic pain. The book also traces the history of pain and pain treatment and indicts the US health care system and the modern pharmaceutical industry for the roles both have played in the ongoing opioid epidemic and the persistence of racial and gender inequities in pain treatment.

Warraich writes vividly and well, using diverse sources that range from Tolstoy (he uses *The Death of Ivan Ilyich* to illustrate the experience of suffering) to unpublished court documents exposing the machinations of the Sackler family. His opening discussion of the physiology of pain is clear and accessible, and he carefully and incisively analyzes the persistence of racial and gender differences in medical imagery and treatment of pain. Histories of discrimination may intensify pain experience, he explains, yet white physicians often "don't trust" pain levels reported by patients of color.

The book's discussion of the long-term memory–forming protein PKMzeta and its potential role in chronic pain sensitization is thought-provoking, drawn from recent research, and a topic that may be new to many readers. When PKMzeta is blocked in mice, chronic pain behaviors fail to develop, suggesting a failure to establish long-term memories of pain-related events and emotions. Further integration of these research findings with other recent work by neuroscientists would have improved this section.

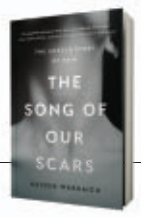
Warraich's tour de force is a fine introduction for someone unfamiliar with this topic, but the narrative presented is not necessarily an "untold story," as the book's subtitle promises. Readers familiar with recent works such as Joanna Bourke's *The Story of Pain* (*1*) or Sam Quinones's *Dreamland* (*2*) will find very little new information here.

The reviewer is at the UCLA Center for Social Medicine and Humanities, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA 90095, USA. Email: meldrum@history.ucla.edu

Like science and medicine, politics and power have long informed our understanding of pain.

PHOTO: ASIAVISION/ISTOCK.COM

The Song of Our Scars: The Untold Story of Pain
Haider Warraich
Basic Books, 2022.
320 pp.



In addition, Warraich's valiant endeavor to integrate the various intersections of pain with science, history, medicine, and sociology is impressive, but he too often sacrifices depth for breadth. His discussion of the early years of US opioid regulation, for example, seems to be entirely drawn from a single article in *Smithsonian* magazine (*3*), and he devotes three paragraphs to the early anti-opium crusader Hamilton Wright but fails to mention the crucial Harrison Act and its impact.

Meanwhile, in chapter 6, Warraich leads into the practice of direct-to-consumer drug advertising, which was first allowed in the US by the Food and Drug Administration in the 1980s, from a discussion of the misleading and hyperbolic direct-to-physician marketing of Valium of the 1960s and 1970s without fully explaining the development and ramifications of this practice. At other points, he digresses briefly into topics that are not uninteresting—discussing itch, for example, in chapter 9—but interrupt the flow of the narrative.

The book's final chapter is the most original and compelling. Here, Warraich shares details of his own recovery from chronic pain through physical therapy and makes a case for alternative treatments, such as hypnotherapy and acceptance and commitment therapy (ACT). He argues for the power of touch and embraces the placebo effect, advocating in particular for the "almost magical quality" an empathetic physician or therapist can bring to the healing process. He also strongly indicts corporate medicine for compromising the ability of physicians to provide empathy, instead emphasizing less-time-intensive therapies, whether those be prescriptions or procedures.

It is hard to fault Warraich's passion for this topic or his thoroughness. But a deeper examination of the crucial aspects of the story of pain, along with some judicious editing of the supporting material, would have made *The Song of Our Scars* a stronger book. ■

REFERENCES AND NOTES

1. J. Bourke, *The Story of Pain: From Prayer To Painkillers* (Oxford Univ. Press, 2014).
2. S. Quinones, *Dreamland: The True Tale of America's Opiate Epidemic* (Bloomsbury, 2015).
3. E. Trickey, "Inside the story of America's 19th-century opiate addiction," *Smithsonian*, 4 January 2018.

10.1126/science.abo5020

Edited by **Jennifer Sills**

Retraction

In the Research Article "Direct imaging discovery of a Jovian exoplanet within a triple-star system" (*1*), we identified a gas giant exoplanet in the star system HD 131399. Follow-up observations by another team (*2*) showed that the detection could have been a false positive. The object might instead have been an unusually fast-moving background source, whose motion is coincidentally aligned with HD 131399, causing it to pass our common proper motion tests. We have now obtained additional observations of the system, spanning a longer time period (*3*). These show that the primary star, HD 131399A, has a parallax at least several times greater than the putative exoplanet, indicating that they are at substantially different distances. This confirms that the object is a background source, not an exoplanet associated with HD 131399. We are therefore retracting the Research Article. All authors agree with this retraction.

Kevin Wagner[1]*, Dániel Apai[1,2], Markus Kasper[3], Kaitlin Kratter[1], Melissa McClure[4], Massimo Robberto[5], Jean-Luc Beuzit[6]
[1]Department of Astronomy and Steward Observatory, University of Arizona, Tucson, AZ 85721, USA. [2]Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721, USA. [3]European Southern Observatory, D-85748 Garching, Germany. [4]Leiden University, 2311 EZ Leiden, Netherlands. [5]Space Telescope Science Institute, Baltimore, MD 21218, USA. [6]Laboratoire d'Astrophysique de Marseille, 13013 Marseille, France.
*Corresponding author.
Email: kevinwagner@arizona.edu

REFERENCES AND NOTES

1. K. Wagner *et al.*, *Science* **353**, 673 (2016).
2. E. L. Nielsen *et al.*, *Astron. J.* **154**, 218 (2017).
3. K. Wagner *et al.*, *Astron. J.* **163**, 80 (2022).

10.1126/science.abq1709

Sri Lanka's endemic rhino-horned lizard (*Ceratophora stoddartii*) has been seen in illegal wildlife markets.

PHOTO: THARAKA S. PRIYADARSHANA

Save Sri Lankan wildlife from foreign smugglers

Sri Lanka, a highly biodiverse tropical island, is home to many vulnerable endemic species, which are targeted by wildlife smugglers. The country's strong species protection laws prohibit the exploitation of endemic species (*1*). In January, a Sri Lankan court imposed the country's largest fine yet (8.6 million Sri Lankan rupees) on three Russian nationals who illegally collected hundreds of animal and plant specimens from Sri Lanka (*2*). However, once smugglers transport a species to another country, they are no longer subject to Sri Lankan law. The international community must work with Sri Lanka to protect Sri Lankan species from illegal wildlife trafficking.

Foreign nationals trying to smuggle rare animals and plants is an ongoing problem in Sri Lanka [e.g., (*3*)]. A recent study found that 12 endangered endemic and range-restricted lizards of Sri Lanka are available in the international pet market (*4*). Most smugglers are based in Europe, where species not listed by the Convention on International Trade in Endangered Species can be freely traded and often sell for a very high price (*5*). In some cases, previously unidentified species of flora and fauna from Sri Lanka have first been described by foreigners based on illegally collected and smuggled specimens (*6*).

Sri Lanka cannot put an end to smuggling alone. Strengthening customs is expensive and requires vigilance in ports, airports, and postal services. Because species in high demand are often found in limited locations, improving site policing, management, and public awareness may be a more cost-effective approach. But these strategies will continue to fall short until weak legislation in Europe and around the world is replaced by laws that effectively prevent wildlife trade (*7*).

Tharaka S. Priyadarshana
Asian School of the Environment, Nanyang Technological University, Singapore City, Singapore. Email: tharakas001@ntu.edu.sg

REFERENCES AND NOTES

1. LawNet, Ministry of Justice, Sri Lanka, "Fauna and flora protection (amendment)" (1993); www.lawnet.gov.lk/fauna-and-flora-protection-amendment-3/#:~:text=an%20ordinance%20to%20provide%20for,connected%20therewith%20or%20incidental%20thereto .
2. T. Razeek, "Laws are not strong enough," *Ceylon Today* (2022); https://ceylontoday.lk/news/laws-are-not-strong-enough-dr-gunawardana#:~:text=%E2%80%9CThe%20laws%20are%20not%20harsh.
3. Sri Lanka Customs, "Biodiversity protection detections" (2021); www.customs.gov.lk/category/biodiversity-protection-detections/.
4. J. Janssen, A. De Silva, *Traffic Bull.* **31**, 9 (2019).
5. M. Auliya *et al.*, *Biol. Conserv.* **204**, 103 (2016).
6. A. A. T. Amarasinghe, R. Pethiyagoda, *Taprobanica* **9**, 133 (2020).
7. S. Altherr, "Stolen wildlife—Why the EU needs to tackle smuggling of nationally protected species" (2014); www.prowildlife.de/wp-content/uploads/2016/02/2014_stolen-wildlife-report.pdf.

10.1126/science.abo4994

Scientists' right to speak to the press

In their Editorial "Strengthening scientific integrity" (21 January, p. 247), A. Nelson and J. Lubchenco summarize the principles that the US Office of Science and Technology Policy (OSTP) has articulated in response to the first report produced by the recently created Scientific Integrity Task Force. As representatives of journalists' organizations, we welcome OSTP's recognition that "federal scientists should be able to speak freely about their unclassified research, including to the press." However, Nelson and Lubchenco do not mention that the task force's report falls short of protecting scientists' rights to speak directly to journalists.

Too often, federal scientists and other staff are prohibited from speaking with journalists without clearance from a supervisor or public information officer. Requiring permission to speak not only violates scientists' First Amendment rights (*1*); it also tramples on the public's right to know and contributes to misinformation and distrust in government. Journalists and others have repeatedly warned (*2*) of the harm done by this form of censorship. In July 2021, 25 groups wrote to OSTP asking the agency to renounce such access restrictions (*3*).

Unfortunately, despite an encouraging statement or two, the new report effectively perpetuates the status quo. It cites approvingly the Obama and Biden administrations' policies allowing media access "in coordination with supervisors and public affairs officials" (*4*). These policies are insufficient. Many state and federal agencies, including the Department of Health and Human Services (*5*) and the Environmental Protection Agency (*6*), require all contacts from reporters to be referred to a press office, ostensibly to ensure that journalists get accurate, timely, and complete information. However, funneling all queries through a press office can have the opposite effect. When journalists are shunted to a press office, requests for interviews or simple queries can be delayed from hours to months, stretching beyond the journalists' deadlines.

The delay could be because public information officers are overwhelmed, or it could be because agency managers are suppressing information that is embarrassing or contradictory to agency messaging. Whatever the reason, policies that ban all unmonitored contact between staff and journalists can deprive the public of critical information, including evidence of abuses of science and misconduct. That undercuts scientific integrity.

Kathryn Foxhall[1]*, Haisten Willis[1], Timothy Wheeler[2]
[1]Committee on Freedom of Information, Society of Professional Journalists, Indianapolis, IN 46244, USA. [2]Freedom of Information Task Force, Society of Environmental Journalists, Washington, DC 20006, USA.
*Corresponding author.
Email: kfoxhall@verizon.net

REFERENCES AND NOTES

1. Brechner Center for Freedom of Information, "Protecting sources and whistleblowers: The first amendment and public employees' right to speak to the media" (University of Florida, 2019); https://brechner.org/wp-content/uploads/2019/10/Public-employee-gag-orders-Brechner-issue-brief-as-published-10-7-19.pdf.
2. "Journalists ask White House for commitment to openness," *SPJ News* (2015).
3. "Scientific integrity joint letter to White House task force" (2021); www.sej.org/sites/default/files/scientific-integrity-joint-letter-to-white-house-task-force07262021.pdf.
4. Scientific Integrity Task Force, "Protecting the integrity of government science" (2022), p. 30; www.whitehouse.gov/wp-content/uploads/2022/01/01-22-Protecting_the_Integrity_of_Government_Science.pdf.
5. US Department of Health and Human Services, "Guidelines on the Provision of Information to the News Media" (2022); www.hhs.gov/about/news/news-media-guidelines/index.html.
6. " Biden's EPA officials: Contact between reporters and staff still banned without controls," PR Office Censorship (2021); https://profficecensorship.blogspot.com/search?q=EPA.

COMPETING INTERESTS

K.F. is on the Freedom of Information Committee of the National Press Club and the Right to Know Committee of the Association of Health Care Journalists. H.W. is a reporter for the *Washington Examiner*. T.W. is associate editor and senior writer for the *Bay Journal*, a nonprofit news organization that covers environmental topics in the Chesapeake Bay watershed.

10.1126/science.abo6353

Let's not abandon Russian scientists

Russian President Vladimir Putin's brutal, unprovoked war against Ukraine deserves all the opprobrium it has received around the world. This abomination warrants every appropriately sized and targeted sanction against the Putin regime that the horrified world can devise. Sizing and targeting sanctions to ensure that their impact on Putin and his designs exceeds the downsides for Western and global interests requires some reflection, however. Such reflection should not be pushed aside in the understandable heat and anguish of the moment. As Western scientists who have engaged in international collaborations in many forms on many topics, we welcome the outpouring of support across the West for Ukrainian scientists, including the introduction of protected visa status for Ukrainians (*1*). At the same time, we urge that our nations' policy-makers and our science and technology communities avoid shunning all Russian scientists for the actions of the Russian government.

Nearly all government-to-government collaboration is understandably on hold now, but not all engagement with Russian scientists should be. Shutting down all interaction with Russian scientists would be a serious setback to a variety of Western and global interests and values, which include making rapid progress on global challenges related to science and technology, maintaining nonideological lines of communication across national boundaries, and opposing ideological stereotyping and indiscriminate persecution.

Many thousands of Russian academics and students live and work in the West. Many of them have criticized the Russian government in the media or have signed widely circulated statements by Russian academics and intellectuals denouncing the Russian invasion [e.g., (*2*)]. Surely these Russians should not be lumped together with leaders of the Russian state. Rather, humanitarian provision should be made to ensure that, as their visas and passports expire, they are not forcibly repatriated to face not only isolation from their Western colleagues but also, very possibly, persecution.

Decisions made in Western countries today about how to deal with Russia and Russians may be in place for a long time and, ultimately, difficult to reverse. We fervently hope that all future decisions about Russian scientists and Russian academic institutions will reflect a balanced appraisal that weighs the likely effectiveness of the measures under consideration in punishing or deterring the Russian state against the undesired adverse impacts on Western and global interests and values.

John Holdren[1]*, Nina Fedoroff[2], Neal Lane[3], Nick Talbot[4], Toby Spribille[5]
[1]Harvard University, Cambridge, MA 02138, USA. [2]The Pennsylvania State University, University Park, PA 16802, USA. [3]Rice University, Houston, TX 77005, USA. [4]The Sainsbury Laboratory, Norwich, Norfolk, UK. [5]University of Alberta, Edmonton, AB T6G 2R3, Canada.
*Corresponding author.
Email: john_holdren@hks.harvard.edu

REFERENCES AND NOTES

1. US Department of Homeland Security, "Secretary Mayorkas designates Ukraine for temporary protected status for 18 months" (2022); www.dhs.gov/news/2022/03/03/secretary-mayorkas-designates-ukraine-temporary-protected-status-18-months.

2. "4,000+ Russian scientists, science journalists pen open letter against Ukraine war," *Science the Wire* (2022); https://science.thewire.in/the-sciences/4750-russian-scientists-science-journalists-sign-open-letter-against-ukraine-war/.

COMPETING INTERESTS

J.H. is a pro-bono senior adviser to the president of the Woodwell Climate Research Center. The center has cooperative relationships with Russian scientists working on monitoring environmental change in the Arctic.

Published online 24 March 2022

10.1126/science.abq1025

TECHNICAL COMMENT ABSTRACTS

Comment on "Impact of neurodegenerative diseases on human adult hippocampal neurogenesis"

Jon I. Arellano, Alvaro Duque, Pasko Rakic

Terreros-Roncal *et al*. (Research Article, 26 November 2021, p. 1106) report the presence of abundant neurogenesis in the adult human hippocampus based mainly on immunolabeling with doublecortin, while identifying very low numbers of progenitors, which are not shown to be proliferative. We discuss this and other flaws in the interpretation of their data that raise questions about their conclusions.

Full text: dx.doi.org/10.1126/science.abn7083

Response to Comment on "Impact of neurodegenerative diseases on human adult hippocampal neurogenesis"

J. Terreros-Roncal, E. P. Moreno-Jiménez, M. Flor-García, C. B. Rodríguez-Moreno, M. F. Trinchero, B. Márquez-Valadez, F. Cafini, A. Rábano, M. Llorens-Martín

Rakic and colleagues challenge the use of extensively validated adult hippocampal neurogenesis (AHN) markers and postulate an alternative interpretation of some of the data included in our study. In Terreros-Roncal *et al*., reconstruction of the main stages encompassed by human AHN revealed enhanced vulnerability of this phenomenon to neurodegenerative diseases. Here, we clarify points and ambiguities raised by these authors.

Full text: dx.doi.org/10.1126/science.abn7270

Comment on "Impact of neurodegenerative diseases on human adult hippocampal neurogenesis"

Arturo Alvarez-Buylla, Arantxa Cebrian-Silla, Shawn F. Sorrells, Marcos Assis Nascimento, Mercedes F. Paredes, Jose Manuel Garcia-Verdugo, Zhengang Yang, Eric J. Huang

Terreros-Roncal *et al*. (Research Article, 26 November 2021, p. 1106) investigated the impacts of human neurodegeneration on immunostainings assumed to be associated with neurogenesis. However, the study provides no evidence that putative proliferating cells are linked to neurogenesis, that multipolar nestin$^+$ astrocytes are progenitors, or that mature-looking doublecortin$^+$ neurons are adult-born. Their histology-marker expression differs from what is observed in species where adult hippocampal neurogenesis is well documented.

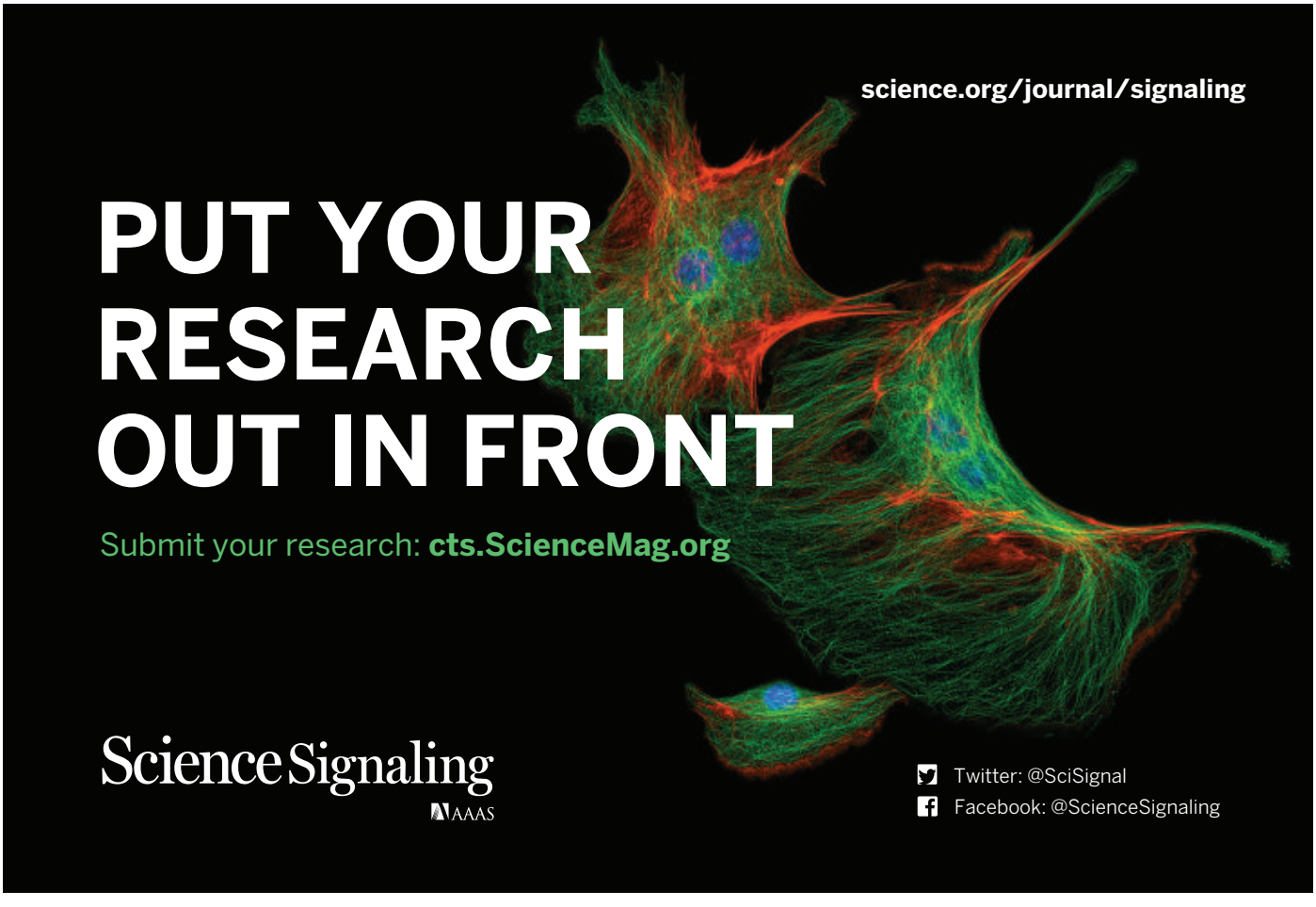
Full text: dx.doi.org/10.1126/science.abn8861

Response to Comment on "Impact of neurodegenerative diseases on human adult hippocampal neurogenesis"

J. Terreros-Roncal, E. P. Moreno-Jiménez, M. Flor-García, C. B. Rodríguez-Moreno, M. F. Trinchero, B. Márquez-Valadez, F. Cafini, A. Rábano, M. Llorens-Martín

Alvarez-Buylla *et al*. provide an alternative interpretation of some of the data included in our manuscript and question whether well-validated markers of adult hippocampal neurogenesis (AHN) are related to this phenomenon in our study. In Terreros-Roncal *et al*., reconstruction of the main stages of human AHN revealed its enhanced vulnerability to neurodegeneration. Here, we clarify ambiguities raised by these authors.

Full text: dx.doi.org/10.1126/science.abo0920



RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

PHOTOS (LEFT TO RIGHT): STUART *ET AL.*; NA *ET AL.*

Mural fragment from a Late Preclassical Maya site with a hieroglyph corresponding to a calendar date

ARCHAEOLOGY

Dating an ancient Maya calendar

A 260-day divinatory calendar known as the *tzolkin* was used across much of Mesoamerica in antiquity, but dating its origins has proven to be difficult. Stuart *et al.* report on the discovery of a painted mural fragment bearing the hieroglyphic date "7 Deer," a day in the calendric cycle, at the Late Preclassic period Maya site of San Bartolo, Guatemala. The fragment with the glyph, along with 10 other examples of text on other fragments, was obtained from a secure archaeological context dated between 300 and 200 BCE. This is the earliest evidence of the 260-day calendar known from the Maya region and is at least 150 years older than previous estimates of its age. —MSA *Sci. Adv.* 10.1126/sciadv.abl9290 (2022).

HYDROGELS

Wrap it up

Conventional stimuli-responsive hydrogel actuators generally suffer from weak actuation force and slow response speed because of the osmotic-driven actuation mechanism. They are also limited in how much pressure they can endure and will collapse or shatter if pushed too hard. Na *et al*. significantly increased the actuation stress of a hydrogel by wrapping the gel in a relatively stiff but flexible semipermeable membrane, which confined the transverse deformation (see the Perspective by Jiang and Song). This effect is similar to the turgor pressure seen in biological cells. The actuation speed can also be enhanced by adding the electrolyte into the water solution and applying an electric field, which reduces the actuation time from hours to minutes. —MSL

Science, abm7862, this issue p. 301

see also abo4603, p. 245

3D PRINTING

Fabulous fabrication

Glass is an important material for micro-optics, microfluidics, and other applications. Fine-scale features and good transparency are often required. Toombs *et al*. combined microscale computed axial lithography with a photopolymer-silica nanocomposite to synthesize fine glass parts. They were able to create optical components, truss and lattice structures, and three-dimensional microfluidics structures. This method should be flexible enough to provide a wide variety of high-quality glass parts for many different applications. —BG

Science, abm6459, this issue p. 308

ORGANIC CHEMISTRY

Controlled release of amine reactants

The formation of carbon–nitrogen bonds is a key step in the synthesis of numerous pharmaceuticals and related compounds. However, it is often not feasible to use the most direct nitrogen-bearing precursors because they inhibit metal catalysts at high concentration through strong coordination. Ali *et al*. report a creative way around this problem for allylic amination reactions. Most of the nitrogen reactants are protected

Hydrogel actuators with a semipermeable membrane move quickly and withstand loads.

as protonated salts, but the products can steadily deprotonate them a few at a time as the reaction progresses. —JSY

Science, abn8382, this issue p. 276

CELL BIOLOGY

Monitoring epithelial integrity

Epithelia provide protective interfaces between various tissues and their surrounding environment. To function as barriers, epithelia must surveil themselves to prevent breaches that can be caused by either physical damage or the formation of aberrant cells. The mechanism by which an epithelium monitors and maintains its integrity is largely unknown. De Vreede *et al*. describe a signaling system in fruit fly organs in which an apically polarized receptor is spatially compartmentalized away from its ligand that circulates in the basal milieu. Polarity defects common to neoplastic cells or wounds mislocalize the receptor, allowing ligand binding and signal transduction. This work shows how an elegant mechanism that detects and repairs breaches to the epithelial barrier has also been adopted to eliminate oncogenic clones and explains how epithelial defects can be recognized as a damage-associated molecular pattern. —BAP

Science, abl4213, this issue p. 297

NEUROSCIENCE

Independent computations within dendrites

Cortical pyramidal neurons typically have an elaborate dendritic tree that receives and integrates the many synaptic inputs targeting the neuron. An open question is how information is represented in dendrites in vivo. Otor *et al*. investigated synaptic computations in the apical tuft of layer 5 pyramidal neurons in the mouse motor cortex using two-photon calcium imaging, behavioral analysis, and cable modeling. Early-branching layer 5 pyramidal neurons showed marked compartmentalization of dendritic calcium signaling, whereas late-branching pyramidal neurons had synchronous tuft activation. *N*-methyl-D-aspartate spikes and cable properties could explain the varying compartmentalization patterns. Compartmentalized activity between hemi-trees was correlated with behavioral outcome. These results indicate a cell-type-dependent dynamic combinatorial code for motor representation. —PRS

Science, abn1421, this issue p. 267

CANCER

OVs and CARs team up against solid tumors

Chimeric antigen receptor (CAR) T cell therapies have limited efficacy against solid tumors, in part because of the lack of tumor-specific antigens. One way to circumvent this problem is through combination treatment with oncolytic viruses (OVs). Egvin *et al*. loaded dual-specific CAR T cells with vesicular stomatitis virus or reovirus to treat mouse models of solid tumors. Systemic treatment with CAR T cells and OVs extended overall survival by restimulating memory CAR T cells with viral T cell receptor specificity. This strategy warrants further investigation for treating patients with solid tumors. —DLH

Sci. Transl. Med. **14**, eabn2231 (2022).

AEROSOL OPTICS

Higher intensity

Light amplification inside atmospheric aerosol particles could affect their photochemical properties. Corral Arroyo *et al*. report that optical confinement can create spatial structuring of the light intensity inside the particle and thereby cause corresponding variations of photochemical rates. Using a combination of x-ray spectromicroscopic imaging and modeling of single ferric citrate particles, the authors predict that photochemical reactions could be sped up in this way by a factor of two to three in atmospheric aerosol particles. —HJS

Science, abm7915, this issue p. 293

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

Scanning electron micrograph of a *C. elegans* nematode

MICROBIOLOGY

A musky warning signal for nematodes

The earthy smell that comes after a rain shower on a warm day is the product of microbes in soil, a terpene called geosmin. Zaroubi *et al*. show that geosmin serves as a deterrent signal to the model nemotode *Caenorhabditis elegans*, which lives in soil and consumes bacteria. Geosmin and related terpenes are not themselves toxic but may be a signal to the worms not to eat certain actinobacteria that do produce other toxic compounds. —MAF

Appl. Environ. Microbiol. 10.1128/aem.00093-22 (2022).

MEDICINE

Gene therapy promotes healing

The rare skin disease recessive dystrophic epidermolysis bullosa (RDEB) is caused by mutations in the *COL7A1* (collagen type VII a1 chain) gene. This gene normally encodes an important component of the basement membrane that connects the epidermis and dermis. RDEB is thus characterized by skin blistering, fibrosis, and susceptibility to infection and cancer. Gurevich *et al*. developed a topical gene therapy, beremagene geperpavec (B-VEC), for RDEB in which *COL7A1* coding sequences are delivered by inactivated herpes simplex virus type 1 (HSV-1) to the skin. Having established that B-VEC restored COL7A1 protein expression in preclinical models, the authors carried out a phase 1/2 trial in nine RDEB patients. B-VEC or placebo was repeatedly applied to matched wounds, and B-VEC promoted COL7A1 expression and wound closure. There was also evidence of correct epidermis-dermis organization. —GKA

Nat. Med. 10.1038/s41591-022-01737-y (2022).

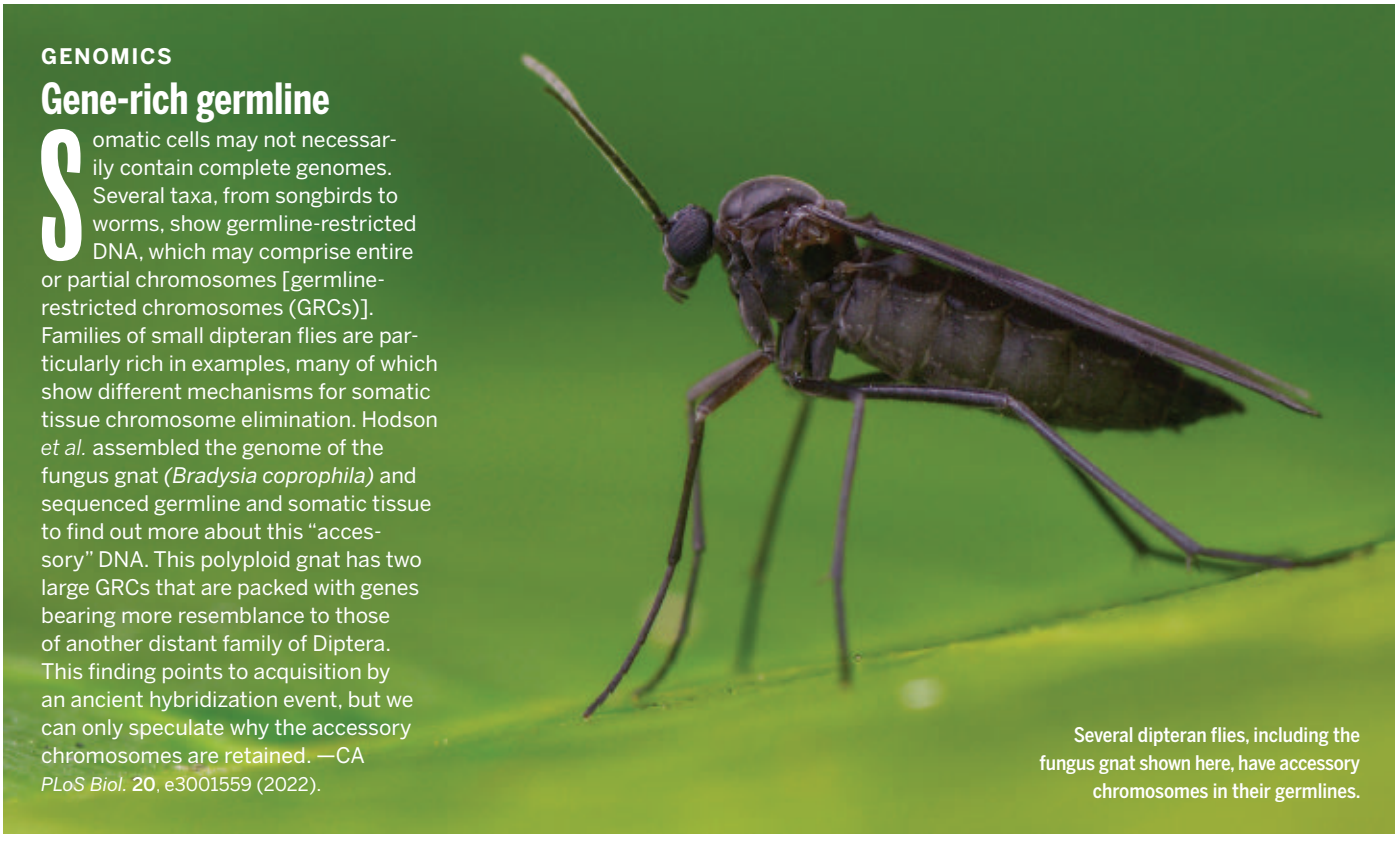
IMAGE: STEVE GSCHMEISSNER/SCIENCE SOURCE

GENOMICS

Gene-rich germline

Somatic cells may not necessarily contain complete genomes. Several taxa, from songbirds to worms, show germline-restricted DNA, which may comprise entire or partial chromosomes [germline-restricted chromosomes (GRCs)]. Families of small dipteran flies are particularly rich in examples, many of which show different mechanisms for somatic tissue chromosome elimination. Hodson *et al.* assembled the genome of the fungus gnat *(Bradysia coprophila)* and sequenced germline and somatic tissue to find out more about this "accessory" DNA. This polyploid gnat has two large GRCs that are packed with genes bearing more resemblance to those of another distant family of Diptera. This finding points to acquisition by an ancient hybridization event, but we can only speculate why the accessory chromosomes are retained. —CA

PLoS Biol. **20**, e3001559 (2022).



Several dipteran flies, including the fungus gnat shown here, have accessory chromosomes in their germlines.

RACIAL DISPARITY

Institutionalized prejudice

Findings that Black drivers in the United States are disproportionately stopped by police officers are well documented. However, understanding what motivates racial disparities in traffic stops has challenged researchers. Such disparities in police activity could arise from individual police officers' racial animus or they may be due to widespread racial bias across the communities in which officers work. Stelter *et al.* connected research on intergroup bias with policing data and examined more than 130 million traffic stops across 1413 US counties. They found that higher levels of racial bias among white Americans at the county level was associated with more traffic stops of Black drivers. These results indicate that in addition to reducing bias among individual police officers, future interventions should account for a larger social context. Prejudice among white residents at the regional level may shape institutionalized practices that motivate police officers to mistreat Black drivers. —EEU

Psychol. Sci. 10.1177/09567976211051272 (2022).

EXTREME PRECIPITATION

Overlook not the weak

Although extreme rainfall typically is thought of as being the result of dramatic storms with intense convection, it also can accompany weak convection events. Xu *et al.* used satellite data from 1998–2014 to show that these weakly convecting extreme precipitation events are as frequent as ones with strong convection globally, each type making up nearly one-third of the total over land. The authors compare and contrast the relative contributions of the two types of extreme rainfall regimes and their geographic distributions and discuss the differences in their dynamic structures and storm environments. These insights should help to improve both our theoretical and observational understanding of extreme rainfall events. —HJS

Geophys. Res. Lett. **49**, 2022GL098048 (2022).

SYNTHETIC BIOLOGY

Plant control of mammalian cells

The clever use of a signaling system borrowed from plants allowed investigators to design a synthetic circuit that controls the size of a population of mammalian cells. Ma *et al.* put elements of the signaling system for the plant hormone auxin into mammalian cells so that they would produce and detect a signal to monitor population size that did not overlap or interfere with endogenous signals. The authors also designed the circuit to catch and destroy "cheater" cells that developed mutations, allowing them to escape regulation. Such population control is an important goal for synthetic biology applications aimed at potential therapeutic applications in humans. —LBR

Cell **185**, 967 (2022).

SUPERCONDUCTIVITY

An unexpected asymmetry

In the physics of cuprate superconductors, the compounds with the highest transition temperatures ("optimal") and those with less than optimal doping ("underdoped") tend to have the most exotic properties. "Overdoped" cuprates, by contrast, are thought to be more conventional. Zou *et al.* used scanning tunneling microscopy to study the tunneling spectra in several overdoped bismuth-based cuprates. The researchers found an asymmetry between the positions of particle and hole coherence peaks, an unusual finding in a superconductor. These findings may have implications for how superconductivity disappears in the transition to a metallic phase. —JS

Nat. Phys. 10.1038/s41567-022-01534-x (2022).

PHOTO: SIMON SHIM/SHUTTERSTOCK

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

ECOSYSTEM ECOLOGY

Declining nitrogen in natural ecosystems

Nitrogen (N) availability is key to the functioning of ecosystems and the cycling of nutrients and energy through the biosphere. However, there is growing evidence that N availability is decreasing in many terrestrial ecosystems. The consequences of declining N availability will be widespread. For example, a decreased concentration of N in leaves reduces the availability of N for insects, contributing to population declines that may then cascade through higher trophic levels. Mason *et al*. reviewed the extent of this phenomenon, and the anthropogenic factors that might be driving it (including climate change and increasing atmospheric carbon dioxide), and discuss how its damaging effects might be mitigated. —AMS

Science, abh3767, this issue p. 261

LIPIDOMICS

Greasing the skin

In multicellular organisms, cells are parts of communities in which an individual contributes to the collective phenotype of the community. Understanding the "social" organization of these cell communities is instrumental to dissecting their physiology and the pathological consequences of their abnormalities. Capolupo *et al*. investigated the lipid metabolism and gene expression of individual human skin cells and found that specific lipid compositions drive cell specialization. Specifically, the authors found that sphingolipids determine the transcriptional programs of fibroblasts populating different layers of the human skin. These results reveal an unexpected role for membrane lipids in the establishment of cell identity and tissue architecture. —LZ

Science, abh1623, this issue p. 262

NEUROIMMUNOLOGY

Give a nod to Nod2 in gut–brain cross-talk

Nod2 is a pattern recognition receptor (PRR) that helps the immune system recognize fragments of bacterial cell walls called muropeptides. Previous work in mice has shown that Nod2 may also play a role in various metabolic and neurologic pathologies. Gabanyi *et al*. report that Nod2 is expressed throughout the brain in reporter mice, including in the hypothalamus (see the Perspective by Adamantidis). When Nod2 was knocked out specifically in inhibitory GABAergic neurons, mice, particularly older female mice, exhibited altered temperature regulation and feeding behavior. Moreover, muropeptides could reach the brain and regulate neurons once there. This work suggests that the brain may sense changes in gut bacteria as a measure of food intake and could serve as the basis for future neurologic and metabolic therapies. —STS

Science, abj3986, this issue p. 263; see also abo7933, p. 248

SIGNAL TRANSDUCTION

mTORC1 as fatty liver disease target

Inappropriate accumulation of fat in the liver causes serious life-threatening disease in humans. In mice, Gosis *et al*. explored the potential beneficial effects of controlling lipid metabolism by preventing signaling by the mTORC1 (mechanistic target of rapamycin complex 1) protein kinase complex (see the Perspective by Ginsberg and Mani). The authors inhibited some, but not all, signaling by mTORC1 by depleting the folliculin protein in the liver. Such mice had increased lipid consumption and decreased lipogenesis and were protected when fed a diet that normally induces nonalcoholic fatty liver disease. These effects appear to result in part from activation of the TFE3 transcription factor. Similar strategies might thus be useful in treating liver disease. —LBR

Science, abf8271, this issue p. 264; see also abp8276, p. 297

MATERIALS SCIENCE

Asymmetrical crystal growth

The shapes of crystalline materials reflect the growth rates of different faces, such as elongation in the fastest-growing direction. In the absence of impurities, boundary walls, or guiding macromolecules, one would expect symmetrical faces to grow at the same rate. Avrahami *et al*. examined crystal growth and arrangement in developing coccoliths, microscopic calcite crystal arrays produced by the unicellular alga *Calcidiscus leptoporus* (see the Perspective by Prywer). The authors found that the {104} faces connected by symmetry relations did not show the same growth rates, thus leading to symmetry breaking in growth and the formation of complex growth habits. This asymmetrical growth is not caused by guide macromolecules, but rather is solely controlled by the diffusion of ions. —MSL

Science, abm1748, this issue p. 312; see also abo2781, p. 240

CATALYSIS

Promoting copper catalysts with C60

Ethylene glycol, a commodity chemical used as a feedstock and antifreeze agent, is synthesized industrially from dimethyl oxalate (DMO) by hydrogenation over precious-metal palladium catalysts at high pressures (typically 20 bars). Copper–chromium catalysts supported on silica as an alternative have required even higher pressures. Zheng *et al*. show that the addition of fullerene (C60) onto copper-silica allows DMO hydrogenation to be performed at ambient pressures with high yield (98%) and without deactivation after 1000 hours (see the Perspective by Gravel and Doris). The use of C60 to stabilize electron-deficient copper species that enhance hydrogen adsorption could likely be applied to other hydrogenation reactions catalyzed by copper. —PDS

Science, abm9257, this issue p. 288; see also abo3155, p. 242

CORONAVIRUS

SARS-CoV-2 spikes the STING pathway

The lungs of patients with severe COVID-19 disease are damaged by inflammation and contain syncytial (or fused) pneumocytes. Liu *et al*. uncovered a mechanism by which severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)–induced cell fusion may contribute to excessive lung inflammation. Cultured epithelial cells expressing the SARS-CoV-2 spike protein fused with cells expressing ACE, the spike protein receptor. The fused cells exhibited DNA damage and micronuclei, which led to the activation of the cytosolic DNA sensor cGAS and the adaptor protein STING and the expression of type I interferon–encoding and interferon-stimulated genes. —JFF

Sci. Signal. **15**, eabg8744 (2022).

INFAMMATORY DISEASE

Troublemakers in the IBD enteric virome

The diverse set of DNA and RNA viruses that comprise the enteric virome are a commonly overlooked component of the complete gut microbiota ecosystem. Adiliaghdam *et al.* compared the composition and immunomodulatory function of the enteric virome from inflammatory bowel disease (IBD)

patients and controls. Virus-like particles enriched from gut tissue or fluid were tested for their immunomodulatory effects on human macrophages or intestinal epithelial cells and introduced into recipient mice. IBD-associated enteric viromes promoted inflammation spontaneously and after dextran sulfate sodium (DSS)–induced colitis, whereas viruses from non-IBD tissue were protective and also suppressed the inflammatory properties of IBD enteric viromes. Viruses distinct to IBD colon tissue included enteroviruses missed in previous fecal virome analyses. These findings illustrate how perturbations in enteric viromes are detected by innate immune system sensors with consequences for the maintenance of normal intestinal homeostasis. —IRW

Sci. Immunol. **7**, eabn6660 (2022).

NATURAL HAZARDS

Homebound seismology

The 2021 Nippes earthquake in Haiti destroyed up to 140,000 homes and killed several thousand people. Despite the large seismic hazard there, Haiti only has a few high-quality seismic stations. Calais *et al*. show that a low-quality seismic network hosted in the homes of volunteers is capable of providing important data for characterizing an earthquake and its aftershocks (see the Perspective by von Hillebrandt-Andrade and Vanacore). The citizen seismic network was particularly important for identifying and determining the likelihood of damaging aftershocks, which is vital information for those responding to the destructive mainshock. —BG

Science, abn1045, this issue p.283; see also abo5378, p. 246

CORONAVIRUS

Say a KIR-full goodbye to autoimmunity

Ly49$^+$CD8$^+$ T cells are a subset of CD8$^+$ T cells that show immunoregulatory activity in mice. Li *et al*. report the existence of a similar CD8$^+$ T cell subset in humans that expresses killer cell immunoglobulin-like receptors (KIRs), a functional parallel of the mouse Ly49 family (see the Perspective by Levescot and Cerf-Bensussan). These cells, which can suppress self-reactive CD4$^+$ T cells, were more abundant in patients with autoimmune conditions such as celiac disease, multiple sclerosis, and lupus, as well as in patients infected with influenza virus or severe acute respiratory syndrome coronavirus 2. When mice selectively deficient in Ly49$^+$CD8$^+$ T cells were infected with viruses, they showed normal antiviral immune responses but eventually developed symptoms of autoimmune disease. KIR$^+$CD8$^+$ T cells may therefore be an important therapeutic target for the control of autoimmune diseases such as "long COVID" that emerge after viral infections. —STS

Science, abi9591, this issue p. 265; see also abp8243, p. 243

NEUROSCIENCE

Soma and dendrite plasticity uncoupled

Although we have detailed knowledge of synaptic and dendritic plasticity in vitro, learning-induced changes in vivo are mostly investigated through unit recordings or imaging of somatic calcium activity. However, we know much less about the functional and plastic properties of dendrites in vivo. D'Aquin *et al*. combined deep-brain imaging with high-resolution, subcellular two-photon microscopy to study the activity of dendrites and somas of identified amygdala neurons over days in awake mice undergoing classical fear conditioning. Sensory stimulation induced compartmentalized dendritic responses that were controlled by dendrite-targeting, somatostatin-positive interneurons. Fear conditioning–induced plasticity was uncoupled between soma and dendrites, possibly reflecting compartment-specific synaptic- and microcircuit-level mechanisms that may increase the computational capacity of amygdala circuits. —PRS

Science, abf7052, this issue p. 266

REVIEW SUMMARY

ECOSYSTEM ECOLOGY

Evidence, causes, and consequences of declining nitrogen availability in terrestrial ecosystems

Rachel E. Mason*, Joseph M. Craine, Nina K. Lany, Mathieu Jonard, Scott V. Ollinger, Peter M. Groffman, Robinson W. Fulweiler, Jay Angerer, Quentin D. Read, Peter B. Reich, Pamela H. Templer, Andrew J. Elmore*

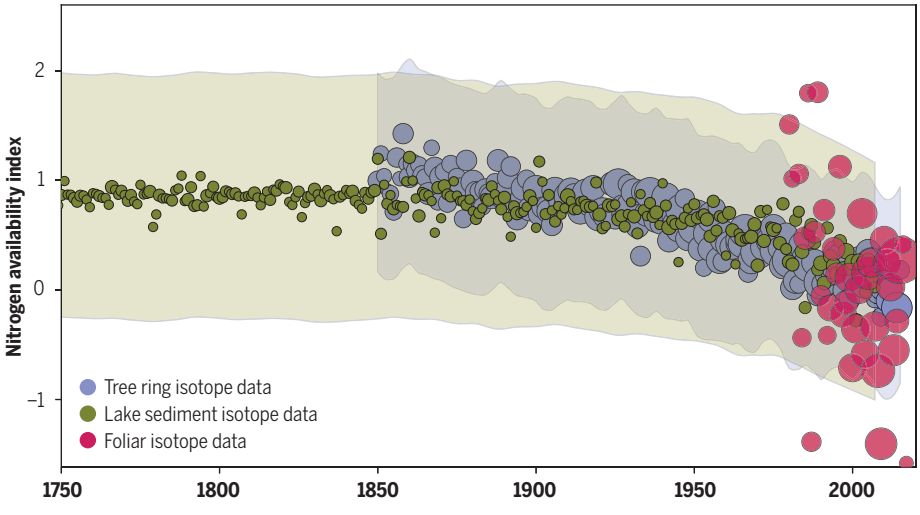
BACKGROUND: The availability of nitrogen (N) to plants and microbes has a major influence on the structure and function of ecosystems. Because N is an essential component of plant proteins, low N availability constrains the growth of plants and herbivores. To increase N availability, humans apply large amounts of fertilizer to agricultural systems. Losses from these systems, combined with atmospheric deposition of fossil fuel combustion products, introduce copious quantities of reactive N into ecosystems. The negative consequences of these anthropogenic N inputs—such as ecosystem eutrophication and reductions in terrestrial and aquatic biodiversity—are well documented. Yet although N availability is increasing in many locations, reactive N inputs are not evenly distributed globally. Furthermore, experiments and theory also suggest that global change factors such as elevated atmospheric CO_2, rising temperatures, and altered precipitation and disturbance regimes can reduce the availability of N to plants and microbes in many terrestrial ecosystems. This can occur through increases in biotic demand for N or reductions in its supply to organisms. Reductions in N availability can be observed via several metrics, including lowered nitrogen concentrations ([N]) and isotope ratios ($\delta^{15}N$) in plant tissue, reduced rates of N mineralization, and reduced terrestrial N export to aquatic systems. However, a comprehensive synthesis of N availability metrics, outside of experimental settings and capable of revealing large-scale trends, has not yet been carried out.

ADVANCES: A growing body of observations confirms that N availability is declining in many nonagricultural ecosystems worldwide. Studies have demonstrated declining wood $\delta^{15}N$ in forests across the continental US, declining foliar [N] in European forests, declining foliar [N] and $\delta^{15}N$ in North American grasslands, and declining [N] in pollen from the US and southern Canada. This evidence is consistent with observed global-scale declines in foliar $\delta^{15}N$ and [N] since 1980. Long-term monitoring of soil-based N availability indicators in unmanipulated systems is rare. However, forest studies in the northeast US have demonstrated decades-long decreases in soil N cycling and N exports to air and water, even in the face of elevated atmospheric N deposition. Collectively, these studies suggest a sustained decline in N availability across a range of terrestrial ecosystems, dating at least as far back as the early 20th century.

Elevated atmospheric CO_2 levels are likely a main driver of declines in N availability. Terrestrial plants are now uniformly exposed to ~50% more of this essential resource than they were just 150 years ago, and experimentally exposing plants to elevated CO_2 often reduces foliar [N] as well as plant-available soil N. In addition, globally-rising temperatures may raise soil N supply in some systems but may also increase N losses and lead to lower foliar [N]. Changes in other ecosystem drivers—such as local climate patterns, N deposition rates, and disturbance regimes—individually affect smaller areas but may have important cumulative effects on global N availability.

OUTLOOK: Given the importance of N to ecosystem functioning, a decline in available N is likely to have far-reaching consequences. Reduced N availability likely constrains the response of plants to elevated CO_2 and the ability of ecosystems to sequester carbon. Because herbivore growth and reproduction scale with protein intake, declining foliar [N] may be contributing to widely reported declines in insect populations and may be negatively affecting the growth of grazing livestock and herbivorous wild mammals.

Spatial and temporal patterns in N availability are not yet fully understood, particularly outside of Europe and North America. Developments in remote sensing, accompanied by additional historical reconstructions of N availability from tree rings, herbarium specimens, and sediments, will show how N availability trajectories vary among ecosystems. Such assessment and monitoring efforts need to be complemented by further experimental and theoretical investigations into the causes of declining N availability, its implications for global carbon sequestration, and how its effects propagate through food webs. Responses will need to involve reducing N demand via lowering atmospheric CO_2 concentrations, and/or increasing N supply. Successfully mitigating and adapting to declining N availability will require a broader understanding that this phenomenon is occurring alongside the more widely recognized issue of anthropogenic eutrophication. ■



Intercalibration of isotopic records from leaves, tree rings, and lake sediments suggests that N availability in many terrestrial ecosystems has steadily declined since the beginning of the industrial era. Reductions in N availability may affect many aspects of ecosystem functioning, including carbon sequestration and herbivore nutrition. Shaded areas indicate 80% prediction intervals; marker size is proportional to the number of measurements in each annual mean.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: rachel.e.mason1@gmail.com (R.E.M.); aelmore@umces.edu (A.J.E.)
Cite this article as R. E. Mason *et al.*, *Science* 376, eabh3767 (2022). DOI: 10.1126/science.abh3767

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abh3767

ISOTOPE DATA: (TREE RING) K. K. MCLAUCHLAN *ET AL.*, *SCI. REP.* **7**, 7856 (2017); (LAKE SEDIMENT) G. W. HOLTGRIEVE *ET AL.*, *SCIENCE* **334**, 1545–1548 (2011); (FOLIAR) J. M. CRAINE *ET AL.*, *NAT. ECOL. EVOL.* **2**, 1735–1744 (2018)

REVIEW

ECOSYSTEM ECOLOGY

Evidence, causes, and consequences of declining nitrogen availability in terrestrial ecosystems

Rachel E. Mason[1]*†, Joseph M. Craine[2], Nina K. Lany[3], Mathieu Jonard[4], Scott V. Ollinger[5], Peter M. Groffman[6,7], Robinson W. Fulweiler[8,9], Jay Angerer[10], Quentin D. Read[1]‡, Peter B. Reich[11,12,13], Pamela H. Templer[9], Andrew J. Elmore[1,14]*

The productivity of ecosystems and their capacity to support life depends on access to reactive nitrogen (N). Over the past century, humans have more than doubled the global supply of reactive N through industrial and agricultural activities. However, long-term records demonstrate that N availability is declining in many regions of the world. Reactive N inputs are not evenly distributed, and global changes—including elevated atmospheric carbon dioxide (CO_2) levels and rising temperatures—are affecting ecosystem N supply relative to demand. Declining N availability is constraining primary productivity, contributing to lower leaf N concentrations, and reducing the quality of herbivore diets in many ecosystems. We outline the current state of knowledge about declining N availability and propose actions aimed at characterizing and responding to this emerging challenge.

Human activities have caused extensive changes in climate, land use, ecosystem function, and biogeochemical cycles, including that of nitrogen (N) (*1*). N is a fundamental component of plant proteins, which are necessary to support the growth of plants and the herbivores that feed upon them. Thus, N availability has a strong influence on the structure and function of many ecosystems. The dominant form of N in the biosphere is highly stable N_2 gas, which humans convert into reactive forms of N through fertilizer production and planting of N_2-fixing crops, and as a by-product of fossil fuel combustion. Application of this reactive N to ecosystems, intentionally or via the deposition of airborne NO_3^- and NH_3, increases N availability, defined here as the supply of N to plants and microbes relative to their demand for N (Box 1). As N availability rises, a cascade of effects occurs, including increased plant N concentrations, shifts in above- and belowground species abundance and diversity, and increased N losses to the atmosphere and aquatic ecosystems. The negative consequences of these changes, which present serious threats to environmental quality and the well-being of human communities, have been the subject of extensive research and discussion (*1*).

At the same time, a growing body of evidence suggests that the problem of excess N coexists with a much less widely recognized issue: declining N availability in terrestrial systems that are not subject to high levels of anthropogenic N inputs. Although humans have more than doubled the total global supply of reactive N (*1*), the largest inputs occur in agricultural and urban areas and downstream locations, and levels of atmospheric N deposition vary widely by region and over time. Large areas of Earth's terrestrial surface, including much of Australia, sub-Saharan Africa, parts of Asia and South America, and vast swaths of boreal forest, have not yet been subject to high levels of N deposition. In addition, elevated N deposition in parts of North America, much of Europe, and some regions of Southeast Asia has decreased in recent decades (*2*, *3*). Therefore, many terrestrial ecosystems are potentially susceptible to changes in ecosystem drivers that may reduce the availability of N. These changes include elevated atmospheric CO_2, rising global temperatures, and altered precipitation and disturbance regimes (*4*–*7*).

Box 1. Nitrogen availability is defined as the supply of N relative to demand by plants and microbes. By accounting for demand, this definition differs from one based solely on N supply. Although an increase in N supply can cause N availability to rise, N availability may decline if demand for N increases by more than any increase in N supply.

Declines in terrestrial N availability can be driven by increases in primary productivity that result in N demand outstripping N supply, decreases in external N inputs, decreases in soil N cycling rates, and/or increases in N losses. Experiments and theory predict declines in N availability in many ecosystems under the influence of a number of global change factors (*4*, *5*, *7*–*9*), but a comprehensive synthesis of N availability metrics, capable of revealing large-scale trends, has yet to be carried out. Acknowledging the substantial evidence of excess reactive N in areas of high anthropogenic inputs, our goal for this paper is to present evidence of declines in N availability in forests, grasslands, and other terrestrial ecosystems outside of agricultural and urban locations.

We show how changes in the N cycle can be evaluated, and we review the likely causes of N availability declines. We then assess their potential consequences for ecosystems and society. Finally, we identify the research that is needed in response to this emerging issue. Akin to trends in atmospheric CO_2 or global temperatures, large-scale declines in N availability are likely to present long-term challenges that will require informed management and policy actions in the coming decades.

Tracking the N cycle

Determining large-scale trajectories of N availability requires monitoring of the N cycle. Yet of all global changes caused by human activity, changes in N availability and cycling are among the most challenging to study. Whereas changes in atmospheric CO_2, precipitation, and atmospheric temperature are routinely monitored and reported globally, tracking the N cycle requires drawing inferences from a suite of indicators collected over a range of scales in space and time (Fig. 1). These indicators include metrics of soil microbial activity, plant N assimilation, and ecosystem N inputs and outputs, which must then be assembled to determine trends in N availability at regional or global scales.

Changes in ecosystem N availability can be inferred from measures of N inputs, internal soil N cycling processes, plant N status, and N losses (Fig. 1). In unfertilized ecosystems, reactive forms of N are added via lightning, biological N_2 fixation, rock weathering, and atmospheric N deposition. These reactive forms of N (NO_3^-, NH_4^+, and small organic molecules) are cycled by plants and soil

[1]National Socio-Environmental Synthesis Center, Annapolis, MD, USA. [2]Jonah Ventures, Boulder, CO, USA. [3]Northern Research Station, USDA Forest Service, Durham, NH, USA. [4]Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium. [5]Earth Systems Research Center, University of New Hampshire, Durham, NH, USA. [6]Advanced Science Research Center, The Graduate Center, City University of New York, New York, NY, USA. [7]Cary Institute of Ecosystem Studies, Millbrook, NY, USA. [8]Department of Earth and Environment, Boston University, Boston, MA, USA. [9]Department of Biology, Boston University, Boston, MA, USA. [10]Fort Keogh Livestock and Range Research Laboratory, USDA Agricultural Research Service, Miles City, MT, USA. [11]Department of Forest Resources, University of Minnesota, St. Paul, MN, USA. [12]Institute for Global Change Biology and School for Environment and Sustainability, University of Michigan, Ann Arbor, MI, USA. [13]Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South Wales, Australia. [14]Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, MD, USA.
*Corresponding author. Email: rachel.e.mason1@gmail.com (R.E.M.); aelmore@umces.edu (A.J.E.)
†Present address: Center for Global Discovery and Conservation Science, Arizona State University, Tempe, AZ, USA.
‡Present address: USDA Agricultural Research Service, Southeast Area, Raleigh, NC, USA.

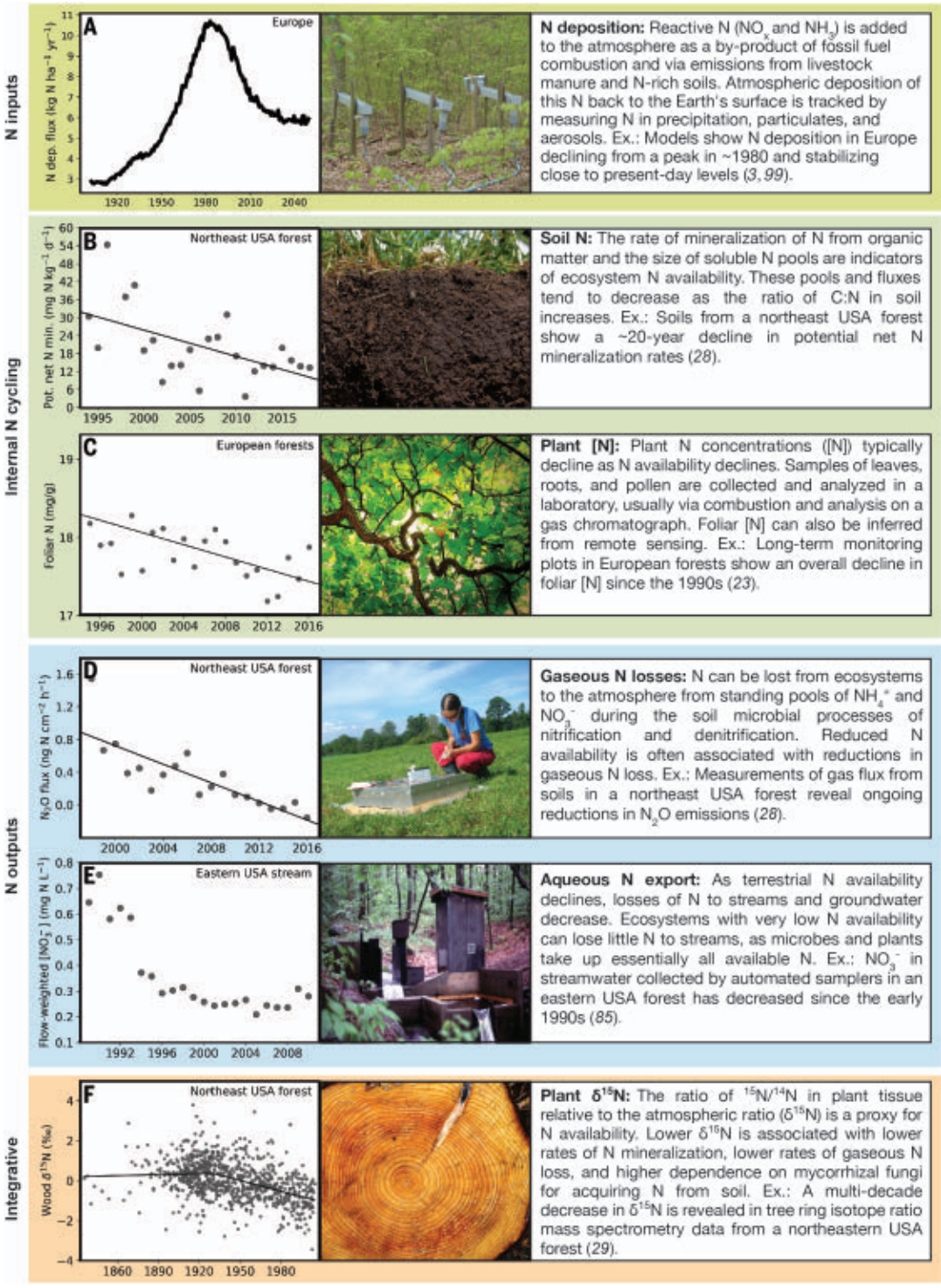

N inputs
A
Europe
N dep. flux (kg N ha⁻¹ yr⁻¹)
1920 1950 1980 2010 2040
N deposition: Reactive N (NO_x and NH_3) is added to the atmosphere as a by-product of fossil fuel combustion and via emissions from livestock manure and N-rich soils. Atmospheric deposition of this N back to the Earth's surface is tracked by measuring N in precipitation, particulates, and aerosols. Ex.: Models show N deposition in Europe declining from a peak in ~1980 and stabilizing close to present-day levels (3, 99).
Internal N cycling
B
Northeast USA forest
Pot. net N min. (mg N kg⁻¹ d⁻¹)
1995 2000 2005 2010 2015
Soil N: The rate of mineralization of N from organic matter and the size of soluble N pools are indicators of ecosystem N availability. These pools and fluxes tend to decrease as the ratio of C:N in soil increases. Ex.: Soils from a northeast USA forest show a ~20-year decline in potential net N mineralization rates (28).
C
European forests
Foliar N (mg/g)
1996 2000 2004 2008 2012 2016
Plant [N]: Plant N concentrations ([N]) typically decline as N availability declines. Samples of leaves, roots, and pollen are collected and analyzed in a laboratory, usually via combustion and analysis on a gas chromatograph. Foliar [N] can also be inferred from remote sensing. Ex.: Long-term monitoring plots in European forests show an overall decline in foliar [N] since the 1990s (23).
N outputs
D
Northeast USA forest
N_2O flux (ng N cm⁻² h⁻¹)
2000 2004 2008 2012 2016
Gaseous N losses: N can be lost from ecosystems to the atmosphere from standing pools of NH_4^+ and NO_3^- during the soil microbial processes of nitrification and denitrification. Reduced N availability is often associated with reductions in gaseous N loss. Ex.: Measurements of gas flux from soils in a northeast USA forest reveal ongoing reductions in N_2O emissions (28).
E
Eastern USA stream
Flow-weighted [NO_3^-] (mg N L⁻¹)
1992 1996 2000 2004 2008
Aqueous N export: As terrestrial N availability declines, losses of N to streams and groundwater decrease. Ecosystems with very low N availability can lose little N to streams, as microbes and plants take up essentially all available N. Ex.: NO_3^- in streamwater collected by automated samplers in an eastern USA forest has decreased since the early 1990s (85).
Integrative
F
Northeast USA forest
Wood δ¹⁵N (‰)
1860 1890 1920 1950 1980
Plant δ¹⁵N: The ratio of $^{15}N/^{14}N$ in plant tissue relative to the atmospheric ratio ($\delta^{15}N$) is a proxy for N availability. Lower $\delta^{15}N$ is associated with lower rates of N mineralization, lower rates of gaseous N loss, and higher dependence on mycorrhizal fungi for acquiring N from soil. Ex.: A multi-decade decrease in $\delta^{15}N$ is revealed in tree ring isotope ratio mass spectrometry data from a northeastern USA forest (29).

Fig. 1. Changes in the N cycle can be detected by monitoring ecosystem N inputs, internal soil N cycling and plant N status, and N losses. In contrast to the well-established monitoring of global atmospheric CO_2, for example, tracking N availability requires observing a comprehensive set of metrics that are often highly variable in space and time and the measurement of which involves considerable effort. Lower N input rates (**A**), smaller pools and fluxes of plant-available soil N (**B**), decreased plant N status (**C**), and lower N losses (**D** and **E**) over time may indicate reductions in N availability, whereas N stable isotope ratios (**F**) provide a measure that integrates over several determinants of N availability. For (A) to (F), respectively, example time series of N availability indicators, here primarily taken from forest ecosystems, are adapted from (*3*, *99*), (*28*), (*23*), (*28*), (*85*), and (*29*). [Photo credits: (A) M. Jonard; (B) Natural Resources Conservation Service/CC BY 2.0; (C) milomingo/CC BY-NC-ND 2.0; (D) A. Contosta; (E) US Forest Service, Northern Research Station; and (F) B. Kasman]

microbes, so measurements of plant N and microbial activity are important tools for tracking N availability. Soil microbes can release N from organic matter into soluble organic N (solubilization), transform organic N to inorganic forms (mineralization), and oxidize NH_4^+ into NO_3^- (nitrification). Soil microbes can also acquire organic and inorganic N from soil solution so that it becomes unavailable to plants in the short term (immobilization). Net mineralization, the balance between mineralization and immobilization, is often estimated from the change in inorganic N in soil solution over a period of time in the absence of plants.

The balance between net mineralization or solubilization and immobilization is highly dependent on factors such as the C:N ratio of organic matter, which is thus an additional indicator of N availability. This ratio is driven by the N concentration ([N]) of plant biomass, which tends to decrease when N availability decreases (*10*). Carbon and N concentrations in samples of plant tissues are measured in the laboratory through combustion and elemental analysis. Transfers of N to water bodies and the atmosphere can also be proxies for N availability, as ecosystem losses of N occur to a greater extent when N availability is high. Quantifying N transfers requires simultaneous measurement of N concentration in water or air and the flux of water or air across the boundary of interest. All of these measurements are important for understanding N cycle changes but are rarely implemented on large spatial or temporal scales, owing to their complexity and cost.

Given the spatial and temporal variability of the N cycle and the number of processes involved, metrics that can integrate N cycling processes into a single value are particularly useful for tracking changes over time. Natural abundance N isotope ratios ($\delta^{15}N$), measured in plant, wood, and sediment samples through mass spectrometry, have emerged as a useful tool for this purpose (*10*). Biological processes that lead to increased N loss via gaseous or leaching pathways tend to discriminate against ^{15}N and favor ^{14}N. Over time, N loss from systems with high N availability—i.e., high N supply relative to demand—therefore increases the $\delta^{15}N$ value of the inorganic N pool that remains available to plants. In addition, as N availability increases, plants rely less on mycorrhizal fungi, which transfer ^{15}N-depleted N to plants (*11*). Consequently, plants growing under conditions of high N availability are enriched in ^{15}N relative to plants growing under low N availability.

Evidence of declining N availability in terrestrial ecosystems

Measurements of $\delta^{15}N$ in leaves, wood, and sediments indicate that declines in N availability extend over a wide geographic area and date back to at least the early 1900s. A global dataset of $\delta^{15}N$ in leaves, composed of ~40,000 measurements from unfertilized locations since 1980, reveals a decrease in N availability throughout the period of record (Fig. 2A) (*12*). Isotopic signatures of recently acquired N are stored in wood, so the $\delta^{15}N$ of wood can also be used to reconstruct extended time series of N availability. Aggregating multiple site-level trajectories in wood $\delta^{15}N$ from forests across the continental US demonstrates a pronounced decrease in N availability since the mid-19th century, particularly in cool, wet regions (Fig. 2B) (*13*). Despite integrating more processes than plant $\delta^{15}N$ over a larger spatial scale, the $\delta^{15}N$ of organic matter in lake

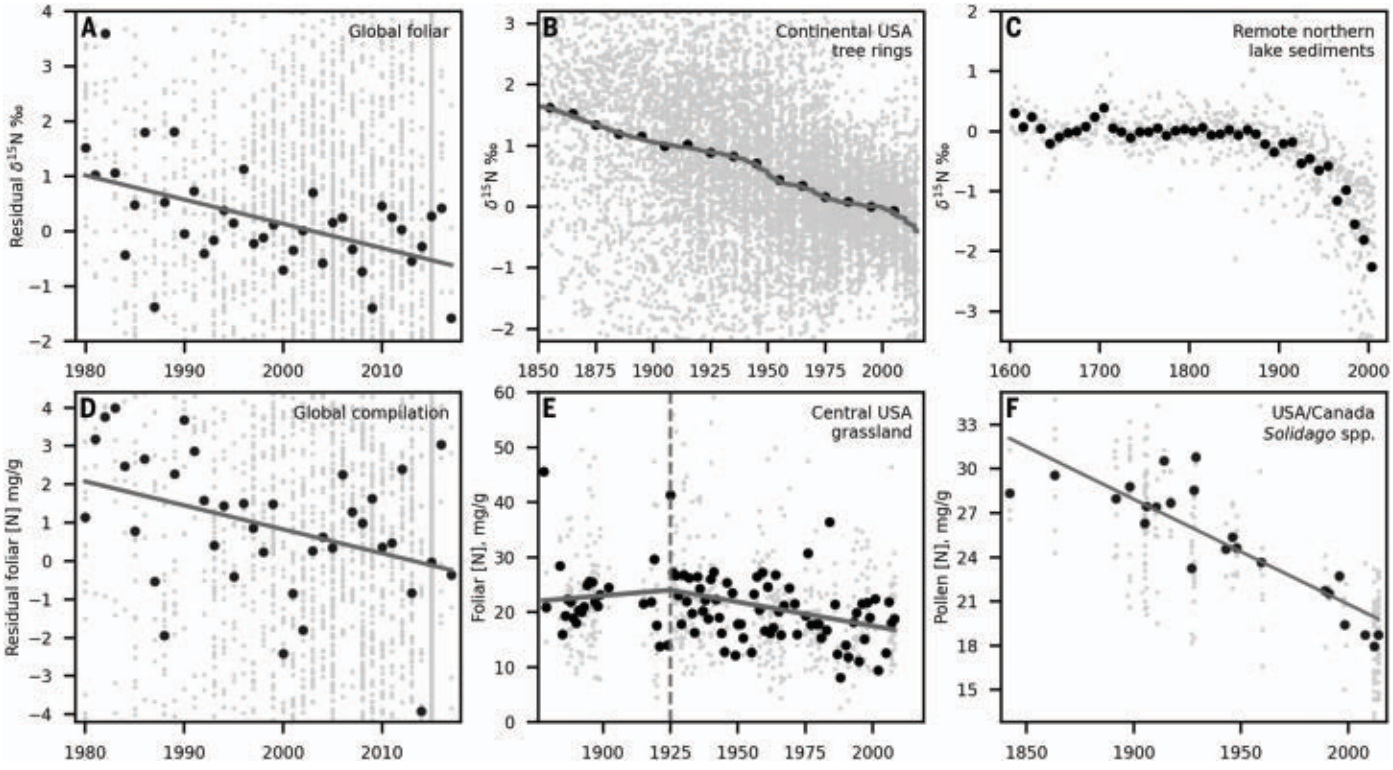


Fig. 2. Evidence of declining N availability comes from long-term global and regional studies. A global foliar $\delta^{15}N$ compilation (**A**) demonstrates a decrease in ecosystem N availability since 1980, whereas tree ring and lake sediment $\delta^{15}N$ datasets (**B** and **C**) from the continental US to the Arctic reveal large-scale declines dating back to at least the early 20th century. Few plant [N] time series cover large temporal and geographic extents. However, statistically significant declines are observed in a global foliar [N] compilation dating back to 1980 (**D**), as well as in long-term records of foliar [N] from a central US grassland (**E**) and pollen [N] from the US and southern Canada (**F**). Data and fits adapted from original publications (*12*, *13*, *15*, *16*, *22*); (C) shows the 25 datasets in (*15*) offset to a common mean. Fits (if any) are as presented in the original papers; all declining trends are significant at the $P < 0.05$ level. Gray points denote individual measurements; black points indicate annual or decadal mean values.

sediments can be used to further extend N availability reconstructions (*14*). Likely owing to the strong influence of local urban and agricultural land use on N availability, no coherent changes in N availability over the past 500 years are apparent at the global scale in lake sediment data (*14*). However, downturns in $\delta^{15}N$ in sediments from remote lakes from the Rocky Mountains to the Arctic (*15*) suggest a decline in N availability over a large area starting around 1895 (Fig. 2C).

The foliar $\delta^{15}N$ record from herbarium samples provides further evidence of long-term declines in N availability. In central and northern US grasslands, the foliar $\delta^{15}N$ of leaves stored in herbaria suggests that N availability has been declining there since roughly 1940 (*16*, *17*). Herbarium studies from Europe and Asia also largely find consistent, declining trends, with data from various species in the western Mediterranean region showing a decrease in foliar $\delta^{15}N$ since the 1920s (*18*) or 1940s (*19*). Foliar $\delta^{15}N$ data from herbarium specimens of *Arabidopsis thaliana* spanning Eurasia and North Africa document a decline in N availability over the period beginning in 1842, although the onset of the decline is not specified (*20*). A set of more recent foliar $\delta^{15}N$ measurements, from the 2000s and 2010s over a ~3000-km transect across the Tibetan Plateau, also exhibit a decline (*21*).

Within individual species, foliar [N] tends to increase with increasing N availability and decrease with decreasing N availability. In parallel to $\delta^{15}N$, a global compilation of foliar [N] measurements since 1980 demonstrates an overall decline (Fig. 2D) (*12*). Herbarium studies show that foliar [N] in grassland species in the central and northern US has decreased by approximately 3 to 8 mg g^{-1} (18 to 30%) since around 1930 (*16*, *17*) (Fig. 2E), and a trend of increasing C:N has been found in *Arabidopsis thaliana* specimens from across this species' broad native range (*20*). Long-term reductions in [N] are not limited to leaves; other herbarium records indicate that [N] in goldenrod (*Solidago* spp.) pollen from multiple locations across the US and southern Canada has decreased by ~10 mg g^{-1} (33%) since the early 1900s (Fig. 2F) (*22*).

Over shorter time scales, ongoing monitoring of European forests demonstrates a general pattern of decreasing foliar [N] (*23*, *24*). Averaged over all species and locations, foliar [N] has been decreasing by 0.04 ± 0.004 mg g^{-1} $year^{-1}$ since at least 1995 (a reduction of 4.4% in 20 years; Fig. 1C) (*23*). Few large-scale foliar [N] time series exist outside of Europe and North America. On the scale of individual sites, comparisons of recent collections and herbarium samples from Panama and the Democratic Republic of the Congo have shown increasing and stable foliar [N], respectively (*25*, *26*). In samples from across China, foliar [N] has increased since the 1980s in tandem with a rise in atmospheric N deposition (*27*).

There are few long-term records that track multiple components of the N cycle, but those that do exist provide valuable insights into the changes occurring as N availability declines. At the Hubbard Brook Experimental Forest (HBEF) in New Hampshire, US, a >50-year monitoring effort covering multiple ecological variables has provided the most detailed published record of declining N availability (*28*). Dendroisotopic and sediment $\delta^{15}N$ records from HBEF imply that the decline in N availability began in approximately 1930, after a period of intense logging (Fig. 1F) (*29*). Export of NO_3^- in streams at HBEF has decreased since the early 1970s, although N deposition at this site began to decrease only in the early 2000s. Gaseous losses of N_2O, a symptom of high N availability in forests, have also declined since measurements began in 1998 (Fig. 1D). Potential net N mineralization and nitrification rates have steadily fallen since the 1970s (Fig. 1B), whereas the C:N ratio of the forest floor has increased (*30*). At other forest sites in the eastern US, long-term monitoring plots reveal trends consistent with declining N availability (*31*), including declines in soil NH_4-N (*32*), forest floor [N] (*33*), and net N mineralization and nitrification (*33*, *34*).

In summary, long-term datasets tracking the N cycle indicate decreasing N availability in multiple locations across Europe and North America, contributing to a pattern of declining N availability in unfertilized terrestrial ecosystems worldwide (*12*). The trend toward lower N availability likely does not extend to locations that receive high levels of anthropogenic N, such as urban and agricultural areas and regions experiencing very high levels of atmospheric N deposition [e.g., China (*27*)], where N availability is characterized by elevated supply. Long-term N availability datasets are scarce in many regions of the world, including most of Asia, the tropics, and the Southern Hemisphere in general. Nonetheless, the forest and grassland ecosystems that exhibit declining N availability represent diverse environments across North America and Eurasia. As well as suggesting that this phenomenon may be affecting large portions of Earth's terrestrial surface, the diverse and widespread nature of the affected ecosystems suggest a shared set of mechanisms underlying the decline in N availability.

Drivers of declining N availability

Multiple environmental changes on both global and local scales may be driving declines in ecosystem N availability (Fig. 3). Elevated atmospheric CO_2 levels (eCO_2) in particular have long been suspected of reducing N availability (*35*). Atmospheric CO_2 has now reached its highest level in millions of years, and terrestrial plants are now uniformly exposed to ~50% more of this essential resource than just 150 years ago. In experiments that expose plants to eCO_2, reduced foliar [N] and increased foliar C:N are consistent outcomes (*4*, *36*). Experiments have also commonly, although not universally, documented N limitation of CO_2-fertilized ecosystems and a reduction in plant-available soil N (*8*, *37*, *38*). Indeed, observational studies find patterns of declining foliar [N] and N availability that are consistent with the expected effects of eCO_2. These patterns include a strong inverse correlation over time between atmospheric CO_2 and plant [N] (*22*, *23*); a spatially uniform decline in foliar [N] and $\delta^{15}N$, suggesting a common driver (*17*); and changes in multiple soil N variables, consistent with the expected consequences of increased C inputs (*28*). Mirroring these outcomes, CO_2 reduction experiments (*8*) and evidence from periods of low atmospheric CO_2 in the planet's history (*39*, *40*) show the reverse effects: increases in plant [N] and N mineralization.

The decrease in foliar [N] under eCO_2 is typically attributed to a set of interlinked processes: increased C assimilation that leads to dilution of foliar N (*36*, *41*), plant responses that reduce investment and incorporation of N into leaves (*41*, *42*), and mechanisms that limit soil N supply (*4*). Support for the dilution hypothesis includes concurrent declines among a suite of foliar nutrients in addition to N (*23*, *24*). Beyond dilution, eCO_2 can lead to changes in N allocation among plant organs, including reductions in RuBisCO (ribulose-1,5-bisphosphate carboxylase-oxygenase) levels in leaves, which in turn increase C assimilation per unit leaf N (*41*, *42*). Reductions in foliar [N] therefore partly imply a decrease in leaf-level N demand. However, this does not necessarily translate to lower N demand at the whole-plant or stand level, as net primary productivity increases with eCO_2. Total plant N uptake may increase along with this growth stimulation (*43*), but not always to the extent necessary to satisfy increased N demand and avoid declines in N availability and foliar [N] (*38*). In addition, when eCO_2 does not lead to an increase in productivity, plant N acquisition appears to be diminished (*4*).

Reductions in plant [N] lead to changes in plant litter chemistry that may influence soil N supplies over time. Elevated C:N in leaf litter, along with an increased flow of C to soil in litter, roots, and root exudates, can promote N immobilization by microbes, reducing the supply of N to plants and potentially further decreasing plant [N] (*35*, *44*, *45*). A decrease in plant-available soil N, both in absolute terms and relative to demand, has been observed in numerous eCO_2 studies (*8*, *37*, *38*, *46*), although other factors such as warming-induced increases

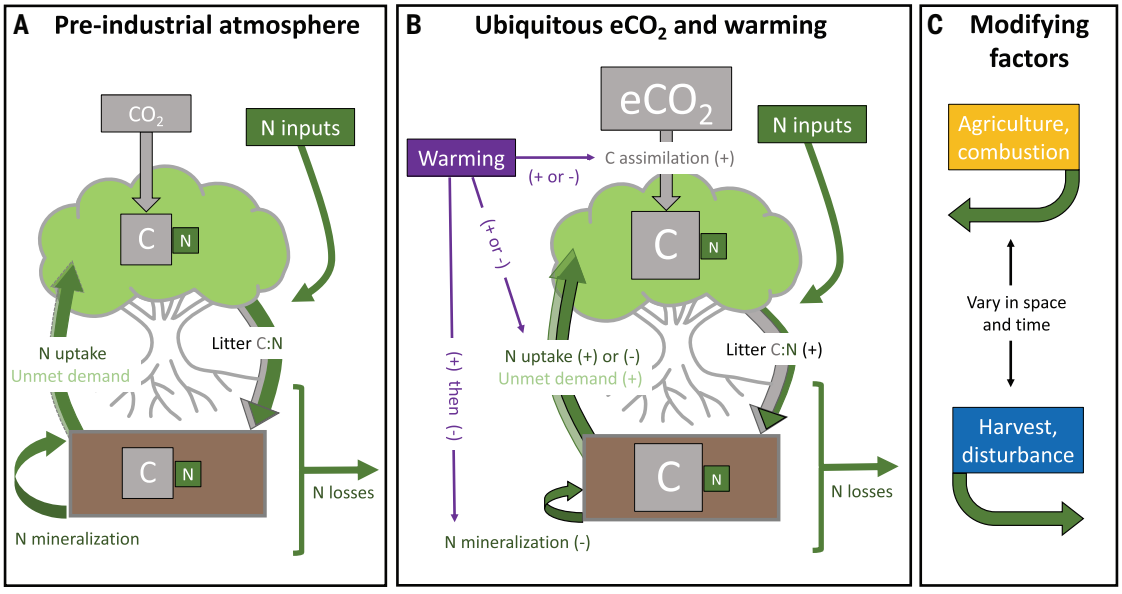


Fig. 3. Multiple global change factors may lead to declines in N availability. In contrast to the atmospheric conditions of pre-industrial times (**A**), present-day eCO_2 (**B**) directly increases assimilation of C by plants, thus increasing foliar C:N and lowering foliar [N]. Plants may invest more in acquiring N from soil but may not be able to obtain sufficient N to meet increased N demand. At the same time, higher C:N in litter may reduce net mineralization of N, lowering soil N supply and plant N uptake and further reducing foliar [N]. Rising temperatures tend to increase N mineralization and plant growth in the short term but may lead to increased N losses and depletion of labile N pools in the longer term. (**C**) Ecosystem N inputs from lightning and biological N fixation are frequently supplemented by inputs from agriculture and combustion, and N outputs can be augmented by harvest of livestock (among other products) and disturbances such as fire. In comparison to eCO_2 and rising global temperatures, these factors vary spatially and can be affected on fairly short time scales by land management and use, air quality regulations, and so forth.

in net mineralization may be able to counteract this reduction (*46*, *47*).

In addition to the direct effects of eCO_2, rising global temperatures affect both plant and microbial processes associated with N supply and demand. Observations across climate gradients demonstrate that plants in warmer environments have lower foliar [N] than those in colder environments (*48*, *49*), suggesting that sustained warming will reduce foliar [N]. Reductions in foliar [N] can result from both long-term (genetic adaptation) and short-term (phenotypic plasticity) processes. Common garden experiments confirm a genetic basis for metabolic adaptation favoring elevated foliar [N] in colder environments (*49*). Ecophysiological studies support the role of warmer temperatures in reducing foliar [N] in conifers grown from seed, showing that short-term metabolic adjustments to warming also reduce foliar [N] (*50*). At the whole-plant scale, warming often improves conditions for growth—one such example is longer growing seasons, which can cause plant N demand to outstrip supply (*51*) and may be associated with reduced plant [N].

Countering plant metabolic adjustments and increases in demand, rising temperatures generally stimulate microbial processes, reducing the residence time of labile organic matter and increasing N supply to plants (*52*). However, with sustained warming, rates of N cycling do not increase for all ecosystems (*53*). Warming can also lead to increases in ecosystem N loss pathways (*6*). Meta-analyses of field warming experiments show mixed results as to whether warming generally increases foliar [N] (*54*, *55*), likely a result of integrating multiple processes related to N supply and demand across diverse ecosystems. Overall, the effects of long-term warming on ecosystem N availability depend on the balance between increases in demand and any increases in supply relative to losses, which will largely be determined by soil organic matter dynamics and any concurrent changes in soil moisture. Especially in dry regions, temperature- and eCO_2-induced changes to soil water deficits could influence both net N mineralization rates and plant N demand.

Elevated atmospheric CO_2 is ubiquitous, and mean annual temperatures are also rising worldwide. Other changes—e.g., in local climate patterns, N deposition rates, and ecosystem disturbance regimes—individually affect smaller areas. Nonetheless, they may have important cumulative effects on global N availability. For example, reduced winter snow cover has been shown to induce soil freezing, fine root damage, reduced net N mineralization, and subsequently reduced N availability (*28*). Similarly, warmer springs can increase vernal asynchrony, lengthening spring conditions conducive to soil microbial mineralization, N leaching, and denitrification, ultimately leading to reductions in N availability during the growing season (*28*, *51*). Projected increases in the frequency and intensity of precipitation may exacerbate N losses through leaching and denitrification (*7*, *56*).

In some areas where N deposition was recently high, air quality regulation has successfully reduced deposition rates. Reductions in N deposition rates tend to result in lower foliar [N] and soil solution NO_3^- (*3*, *34*). Nevertheless, decreases in N deposition cannot fully explain declining terrestrial N availability. In many cases, the decline began before N deposition started to decrease (*28*), dates back to before N deposition became widespread in the 1950s (*13*, *17*, *22*), and/or is taking place in locations in which N deposition has never reached high levels (*17*).

Altered ecosystem disturbance regimes and associated losses of N may also have contributed to historic and ongoing declines in N availability. Through harvesting of biomass, N has been continuously—and is increasingly (*57*, *58*)—exported from ecosystems (in the form of livestock, timber, and other products) and transported to the most-populated watersheds. The frequency of fires is also rising in many locations and is associated with higher N losses over decadal time scales (*5*, *59*): In savanna grasslands and broadleaf forests, frequent burning has been found to reduce

soil N by almost 40% over six decades (*5*). Where frequent N losses occur without substantial inputs, such as in most rangelands that are grazed without use of fertilizers and in situations where supplemental feeding is not feasible (e.g., most pastoral livestock systems), a long-term decline in N availability will be difficult to avoid.

Consequences of declining N availability

Nitrogen availability affects multiple ecosystem processes and services. Researchers have begun to investigate the effects of declining N availability on ecosystem function and have found evidence of impacts on the global C cycle, herbivore nutrition, and water quality. For example, although terrestrial primary productivity has increased globally in response to eCO_2 (*60*) and longer growing seasons (*61*), helping to buffer anthropogenic CO_2 emissions, declining N availability likely constrains this response (*51*, *62*, *63*). In global change experiments, eCO_2 treatment alone tends to increase net primary productivity and C storage less than CO_2 enrichment combined with N addition (*64*), indicating a reduction in N availability under eCO_2 that limits primary productivity. Satellite observations of primary productivity have confirmed the dependence of CO_2 fertilization on N and also suggest that a recent weakening of the CO_2 fertilization effect is due in part to declining N availability (*65*). Without the widening gap between N supply and demand, the terrestrial C sink would likely be greater.

The reductions in foliar [N] that accompany declining N availability may reduce the growth and reproduction of herbivorous insects. Insect herbivore growth rates and abundance are strongly dependent on the availability of protein as a food source (*66*), and as protein and N concentrations are positively correlated in leaves, plant N concentrations are a good index of host plant quality for insect herbivores (*67*). Experiments show that insect herbivores may initially respond to reduced plant [N] by increasing consumption (*68*, *69*), but declines in plant [N] ultimately reduce insect growth, survival, reproduction, and population size (*70*–*72*). At the community level, decreases in plant [N] are expected to change the relative abundance of insect species present (*73*) and reduce biomass transfer to higher trophic levels (*74*).

In a central US grassland ecosystem, a 36% decline in grasshopper abundance has been linked to a 42% decline in foliar [N] that has taken place over the past 30 years (*75*). In eCO_2 experiments, a 16% reduction in foliar [N] (with concomitant changes in other plant characteristics) resulted in a 22% decrease in insect herbivore abundance (*69*). Given the magnitude of foliar [N] decline seen in long-term datasets, declining foliar [N] may be contributing substantially to global declines in terrestrial insect abundance that have averaged ~9% per decade since 1925 (*76*). Although reductions in insect abundance have been attributed to factors as varied as rising temperatures and agricultural practices, reduced N availability could exacerbate the effects of factors such as pesticides (*77*) and provide a unifying explanation of patterns observed across ecosystems.

Because N concentrations are tightly linked among plant organs, decreases in foliar [N] are likely accompanied by decreases in root, stem, and pollen [N] and would thus also affect insects that consume these other parts of the plant. For example, reductions in pollen N concentration (Fig. 2F) (*22*), which can reduce the ability of bees to resist pests and overwinter, could contribute to declines in pollinator abundance. Whereas declining N availability may have negative effects on insects, so too may N enrichment (*78*). This emphasizes the need to better understand the mechanisms linking N availability and insect performance and identify management and policy actions that avoid both excess and insufficient N.

As in insects, growth rates in vertebrate herbivores are often limited by feed protein supply (*79*). Although there are few long-term records of dietary quality for herbivores, regular collection of fecal samples from cattle grazing on rangelands across the US allows the reconstruction of dietary protein concentrations. Independent of any changes in precipitation, crude protein concentrations have been declining since measurements began in 1995 (*80*). Because of changes in genetics, cattle weights in the US have increased over the past 50 years despite the decline in dietary quality (*81*). However, for other large herbivores that do not receive protein supplementation or undergo strong genetic selection by humans, the decline in foliar [N] may be reducing body size and reproduction. For example, bison from regions with lower plant protein concentrations gain weight more slowly, and low protein concentrations are associated with lower reproduction rates (*82*) (Fig. 4).

Terrestrial N availability strongly influences the N loading of headwater streams and, ultimately, coastal receiving waters. In principle, decreasing streamwater inorganic N concentration influences the amount and biochemical composition of primary producers (*83*), initiating bottom-up effects that can propagate to higher trophic levels (*84*). Long-term records from many stream ecosystems lacking substantial anthropogenic N inputs have exhibited declining inorganic N concentrations in recent decades (*28*, *85*, *86*). Under these conditions, primary and secondary production are expected to decline, and recently reported declines in aquatic insect populations (*87*) are consistent with these expectations. In coastal systems, large reductions in N inputs have caused declines in fish productivity (*88*, *89*) and fish landings (*90*). Therefore, changes in inputs from terrestrial ecosystems experiencing declining N availability have the potential to affect watershed N budgets and coastal ecosystem processes.

However, most aquatic systems continue to receive N inputs from local agriculture, wastewater, and other anthropogenic sources. Declining terrestrial N availability in the parts of these watersheds that are not subject to heavy anthropogenic N loading would initially be expected to improve downstream aquatic conditions overall. This may include increases in some stream taxa (*76*) and in coastal ecosystems, greater water clarity, increased abundance of submerged macrophytes, and increases in oxygen concentrations (*91*). Given the continued high level of anthropogenic N inputs to coastal systems, any effects of declining terrestrial N availability will most likely be difficult to detect. Long-term watershed monitoring will be required to correctly associate water quality improvements with improvements in N management and response to external factors such as rising CO_2.

Responses to declining N availability

Since the mid-20th century, increasingly high-profile research and discussion has focused on the negative effects of excess N on terrestrial and aquatic ecosystems. As a result, reduction of anthropogenic N inputs to the Earth system is widely recognized as a high priority. The emerging evidence of a large-scale decline in N availability in unmanaged ecosystems does not contradict previous work that has documented the effects of excess N. Nitrogen is certainly being applied in excess to many agricultural ecosystems, high levels of atmospheric deposition can occur, and the consequences of excess N addition for coastal receiving waters are substantial. Instead, the evidence presented here is a strong indication that the world is now experiencing a dual trajectory in N availability (*12*), in which many areas are exposed to excessive levels of reactive N while others are experiencing declining N availability.

Fundamentally, declining N availability adds to the already overwhelming case for reducing anthropogenic CO_2 emissions. Emissions reductions are needed to stabilize the climate system and moderate ecosystem changes that are a direct consequence of eCO_2. In tandem with much-needed curbs on emissions, research, management, and policy attention to declining N availability should also become a priority (Fig. 5).

Monitoring and assessment

Despite strong indications of declining N availability in many places and contexts, spatial and temporal patterns are not yet well enough understood to efficiently direct global management efforts. A comprehensive assessment

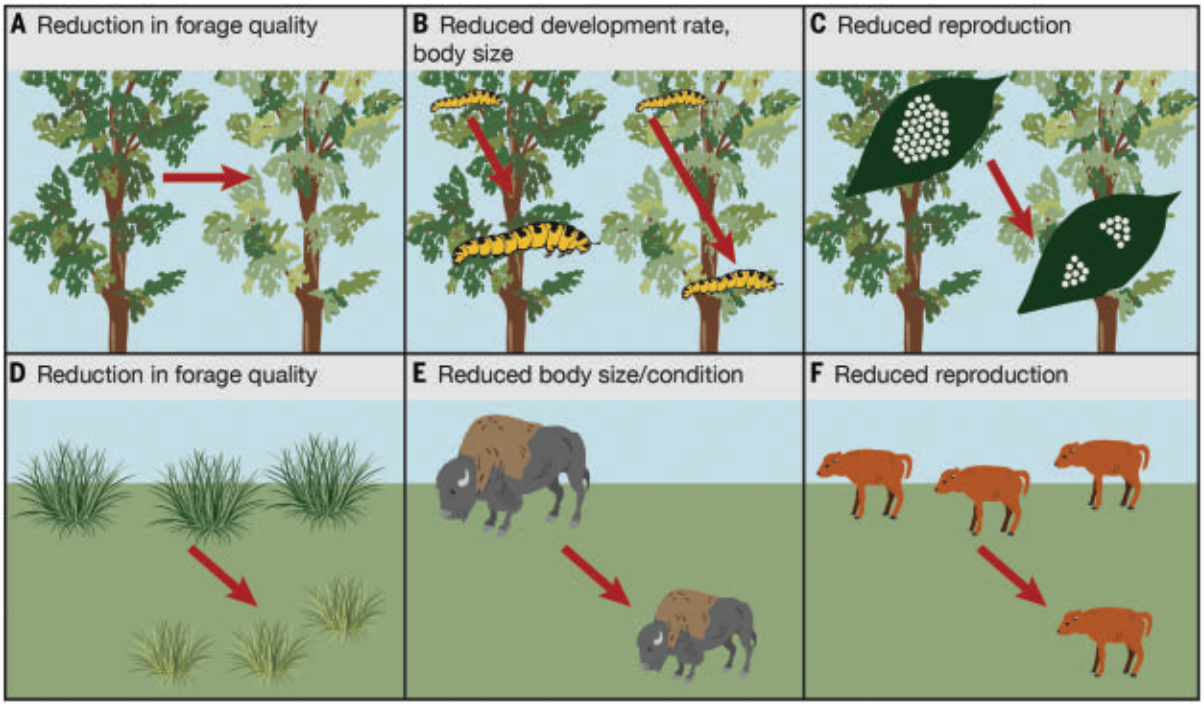

A Reduction in forage quality
B Reduced development rate, body size
C Reduced reproduction
D Reduction in forage quality
E Reduced body size/condition
F Reduced reproduction

Fig. 4. Impacts of declines in foliar N concentrations on herbivore performance. Reduction in forage quality (**A** and **D**) may result in reduced herbivore body size and/or development rate (**B** and **E**) and reproduction (**C** and **F**) because herbivore growth rates and populations are often limited by protein availability.

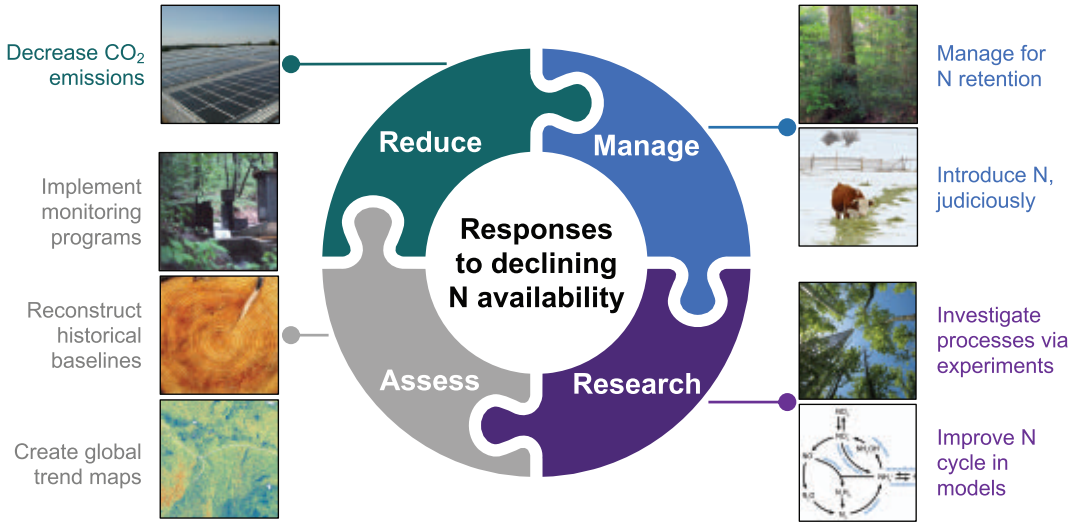


Fig. 5. Addressing the root cause of declining N implies reducing CO_2 emissions, whereas ecosystem assessments and continued research are needed to inform management actions. Variation in N availability trends over space and time can be elucidated through field monitoring campaigns, reconstruction of historical records, and creation of maps via hyperspectral remote sensing techniques. Continued experimentation is required to better understand the processes driving, and resulting from, reductions in N availability. Incorporating this knowledge in ESMs will clarify how declining N may affect the ability of ecosystems to buffer CO_2 emissions. [Image credits (counterclockwise from top left): h080/CC BY-SA 2.0; US Forest Service, Northern Research Station; B. Kasman; S. Ollinger; M. Kirk/CC BY-SA 4.0; US Department of Energy, Oak Ridge National Laboratory/CC BY 2.0; L. Lamsa/CC BY 2.0; and N. Tonelli/CC BY 2.0]

program would involve monitoring of N availability metrics such as N concentrations in plant tissues, net N mineralization in soils, N concentrations in aquatic ecosystems, and herbivore dietary quality. Systematic collection of satellite hyperspectral remote sensing data will soon help facilitate assessment of foliar [N] across broader spatial scales than those possible with current airborne instrumentation or field sampling. Continental-scale monitoring of N deposition has provided important information about N supply, but complex spatial patterns of rising, falling, and stabilizing trends (*2*, *3*) justify expanded instrumentation.

Collectively, these data could be assembled into an annual state-of-the-N-cycle report that would represent a comprehensive resource for scientists, managers, and policy-makers. Data products could include global maps of changing N availability. For example, net primary productivity trends and levels of (or trends in) atmospheric N deposition can be used to estimate N demand and supply and to generate maps of N availability (Fig. 6). Although this captures only a subset of N availability drivers, the resulting visualization is consistent with the evidence of declining N availability in North America and Europe and of rising N availability in China. It also suggests that large areas outside these relatively well-studied regions may be experiencing decreasing N availability (*12*). Additional N availability datasets would allow such maps to be calibrated and refined.

In addition to contemporaneous monitoring, reconstructions of N availability from herbaria, tree rings, and sediments are necessary to understand historical trajectories and set baselines that can be used to guide management efforts. Combining such datasets, similar to how multiple proxies have been intercalibrated to reveal past climates and atmospheric chemistry, would provide additional long-term context for interpreting recent trends. To demonstrate the potential of long-term, multi-proxy N availability reconstructions, we adapted the standard paleoecological approach of intercalibrating different records [δ^{15}N of lake sediments, tree rings, and foliar samples (*12*, *13*, *15*)] to produce a ~250-year record of δ^{15}N spanning continental-to-global scales (Fig. 6). The combined record shows that N availability was fairly constant until a decline began in the early 20th century. Expanding the geographic and temporal coverage of the data used to construct similar diagrams will provide new perspectives on recent trends and help to explain regionally specific causal mechanisms.

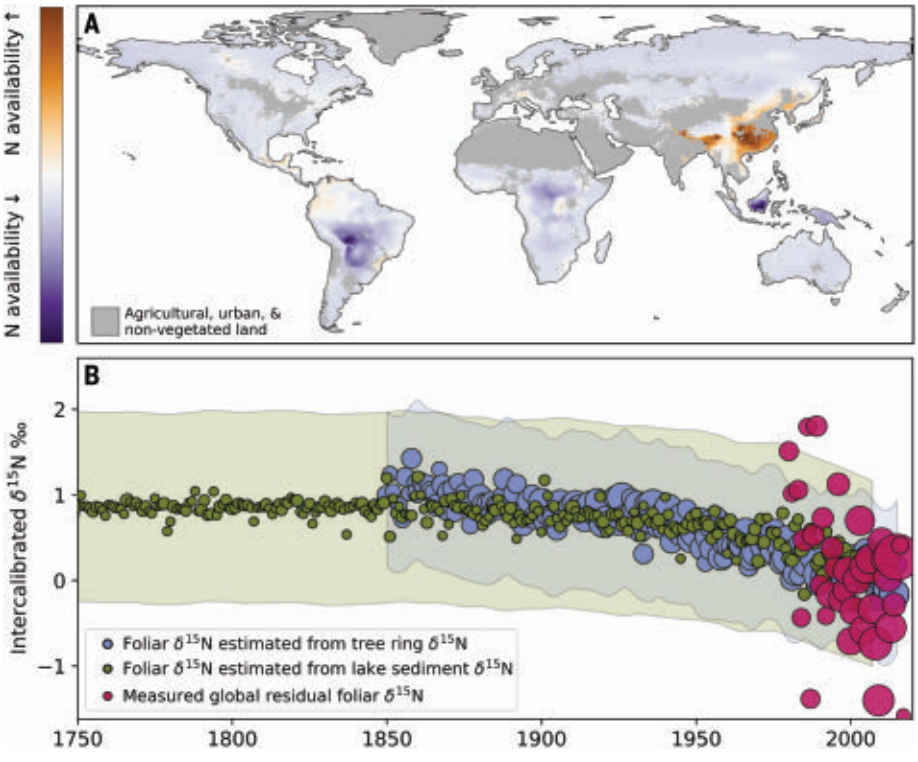


Fig. 6. Mapping the drivers of N supply and demand, and intercalibrating historical N availability records, will provide novel perspectives on trends in global N availability. (**A**) Comparing trends in net primary productivity (NPP) (approximating N demand) and levels of (or trends in) atmospheric N deposition (approximating N supply) suggests increasing N availability in high-deposition regions such as China and declining N availability in many other regions. This visualization was produced by subtracting global maps of N deposition and N deposition trends (*2*) from a global map of trends in NPP (*100*) after normalizing all quantities by dividing by their standard deviation and centering the N deposition map at 10 kg ha^{-1} year^{-1}. Future work that incorporates information about other drivers of N supply and demand will provide a more comprehensive picture of changes in N availability. (**B**) Intercalibrating records from leaves, tree rings, and lake sediments (Fig. 2, A to C) suggests that the declines in N availability began in the early industrial era. Data on tree ring (*13*) and lake sediment (*15*) δ^{15}N from North America were intercalibrated with a global foliar δ^{15}N time series (*12*) using a Bayesian model that included an ARMA (autoregressive moving average) error structure to account for temporal autocorrelation. Shaded areas indicate 80% prediction intervals; marker size is proportional to the square root of the number of measurements included in each annual mean. One foliar δ^{15}N point (at 1982, 3.6) is beyond the scale of the plot.

Research challenges

Beyond monitoring and reconstructions, research into the ecosystem processes involved in declining N availability is needed (*3*, *34*). In ecosystems that have been subject to high levels of anthropogenic N inputs, such as central and western Europe, declining N availability may present welcome opportunities for restoration. However, hysteresis, differential responses of different ecosystem components, and concurrent changes in other environmental conditions complicate predictions of the speed and direction of ecosystem trajectories under declining N inputs (*3*, *34*, *88*).

Declining N availability is also likely to affect plant N:P stoichiometry, which in turn influences plant, herbivore, and microbial community composition (*92*). At the same time, alterations in the availability of other nutrients mean that changes in ecosystem stoichiometry are not entirely predictable. For example, despite declining N deposition, foliar [P] has decreased more rapidly than foliar [N] in European forests, leading to an increase in foliar N:P (*23*, *24*). Although we focus on declines in N availability here, research into the effects of declining N availability will need to consider changes in the availability of other nutrients as well.

Field eCO_2 experiments have provided valuable insights into how declining N availability may arise and progress, but relatively few are operating today, and few have examined multiple global changes (e.g., eCO_2, warming, precipitation change, biodiversity change) simultaneously. Restoring and expanding such studies would improve our understanding of the processes that are the basis of Earth system models (ESMs). Although it is well recognized that N availability is a fundamental constraint on the ability of the biosphere to absorb CO_2 (*62*), only around half of current ESMs include interactions between the C and N cycles. In general, when N cycling is included in ESMs, the projected ability of terrestrial ecosystems to absorb CO_2 emissions tends to decrease because of constraints on CO_2 use due to N limitations (*93*). These models vary in how various components of the N cycle are represented, and they have yet to be parameterized with global N availability datasets.

Further research into the consequences of declining N availability is also needed. The possible role of declining foliar [N] in ongoing declines in insect populations (*75*, *94*) merits particular attention. Declining ecosystem N availability may have relatively direct implications

for human health and well-being that should be investigated. For example, lower protein concentrations in grazing livestock diets may disproportionately affect those who do not have the resources to acquire supplemental feed for their animals. Low-N plants can also increase the abundance of certain locust species, so continued research into feasible and locally appropriate land management practices that promote soil fertility will be valuable (*95*).

Depending on the context, responses to declining N availability may require meeting increased N demand, compensating for N removed in harvested products, reversing declines in plant [N], and promoting C sequestration. Nutrient additions are commonly used to achieve this kind of ecosystem management goal; for example, salmon carcasses and fertilizers have been added to streams to support salmon populations (*96*), and N fertilization is routinely used on improved pastures to increase biomass and enhance forage quality for livestock. Such actions could be implemented at larger scales, but this would be contentious given that fertilizer use has historically led to negative impacts such as eutrophication of aquatic systems.

Moreover, the presence of multiple concurrent environmental changes suggests that further research is needed to design N-addition interventions that achieve the intended effects. For example, decreases in foliar [N] under eCO_2 are partly a consequence of fundamental changes to plant metabolic function in a high-CO_2 environment, and foliar [N] tends to remain depressed in experiments that combine moderate N additions and eCO_2 (*97*). Given that concentrations of P, S, Ca, Mg, and K have decreased in European forests (*23*, *24*), inputs of N alone may not be sufficient to remove nutrient limitations to primary productivity and could induce further nutritional imbalances (*24*). Overall, any N-addition programs will require careful, evidence-based design, with costs, logistical challenges, and implications for water quality (*96*) and greenhouse gas emissions (*98*) taken into account.

Our evolving understanding of the Earth system has led to new concerns about N insufficiency after years of attention to surplus N in the environment. An integrated suite of responses will be needed to simultaneously manage both of these problems. Given the potential implications of declining N availability for food webs, carbon sequestration, and other ecosystem functions and services, it is important that research, management, and policy actions be taken before the consequences of declining N availability become more severe. It can be difficult to create a shared understanding of the N cycle and the many effects of N on ecosystem health and human well-being. The combination of excess N and declining N availability, in which outcomes vary widely across landscapes, adds to this challenge. Developing dialogues among diverse stakeholders—scientists, ecosystem managers, and others—will be necessary for alleviating and adapting to declining N availability in an N-rich world.

REFERENCES AND NOTES

1. J. N. Galloway *et al*., Transformation of the nitrogen cycle: Recent trends, questions, and potential solutions. *Science* **320**, 889–892 (2008). doi: 10.1126/science.1136674; pmid: 18487183
2. D. Ackerman, D. B. Millet, X. Chen, Global Estimates of Inorganic Nitrogen Deposition Across Four Decades. *Global Biogeochem. Cycles* **33**, 100–107 (2019). doi: 10.1029/2018GB005990
3. A. Schmitz *et al*., Responses of forest ecosystems in Europe to decreasing nitrogen deposition. *Environ. Pollut.* **244**, 980–994 (2019). doi: 10.1016/j.envpol.2018.09.101; pmid: 30469293
4. Z. Feng *et al*., Constraints to nitrogen acquisition of terrestrial plants under elevated CO_2. *Glob. Change Biol.* **21**, 3152–3168 (2015). doi: 10.1111/gcb.12938; pmid: 25846203
5. A. F. A. Pellegrini *et al*., Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity. *Nature* **553**, 194–198 (2018). doi: 10.1038/nature24668; pmid: 29227988
6. L. Li *et al*., Terrestrial N_2O emissions and related functional genes under climate change: A global meta-analysis. *Glob. Change Biol.* **26**, 931–943 (2020). doi: 10.1111/gcb.14847; pmid: 31554024
7. H. Ren *et al*., Exacerbated nitrogen limitation ends transient stimulation of grassland productivity by increased precipitation. *Ecol. Monogr.* **87**, 457–469 (2017). doi: 10.1002/ecm.1262
8. R. A. Gill *et al*., Nonlinear grassland responses to past and future atmospheric CO_2. *Nature* **417**, 279–282 (2002). doi: 10.1038/417279a; pmid: 12015601
9. L. E. Street, S. Caldararu, Why are Arctic shrubs becoming more nitrogen limited? *New Phytol.* **233**, 585–587 (2022). pmid: 34820852
10. J. M. Craine *et al*., Global patterns of foliar nitrogen isotopes and their relationships with climate, mycorrhizal fungi, foliar nutrient concentrations, and nitrogen availability. *New Phytol.* **183**, 980–992 (2009). doi: 10.1111/j.1469-8137.2009.02917.x; pmid: 19563444
11. E. A. Hobbie, P. Högberg, Nitrogen isotopes link mycorrhizal fungi and plants to nitrogen dynamics. *New Phytol.* **196**, 367–382 (2012). doi: 10.1111/j.1469-8137.2012.04300.x; pmid: 22963677
12. J. M. Craine *et al*., Isotopic evidence for oligotrophication of terrestrial ecosystems. *Nat. Ecol. Evol.* **2**, 1735–1744 (2018). doi: 10.1038/s41559-018-0694-0; pmid: 30349095
13. K. K. McLauchlan *et al*., Centennial-scale reductions in nitrogen availability in temperate forests of the United States. *Sci. Rep.* **7**, 7856 (2017). doi: 10.1038/s41598-017-08170-z; pmid: 28798386
14. K. K. McLauchlan, J. J. Williams, J. M. Craine, E. S. Jeffers, Changes in global nitrogen cycling during the Holocene epoch. *Nature* **495**, 352–355 (2013). doi: 10.1038/nature11916; pmid: 23518563
15. G. W. Holtgrieve *et al*., A coherent signature of anthropogenic nitrogen deposition to remote watersheds of the Northern Hemisphere. *Science* **334**, 1545–1548 (2011). doi: 10.1126/science.1212267; pmid: 22174250
16. K. K. McLauchlan, C. J. Ferguson, I. E. Wilson, T. W. Ocheltree, J. M. Craine, Thirteen decades of foliar isotopes indicate declining nitrogen availability in central North American grasslands. *New Phytol.* **187**, 1135–1145 (2010). doi: 10.1111/j.1469-8137.2010.03322.x; pmid: 20553396
17. E. N. J. Brookshire, P. C. Stoy, B. Currey, B. Finney, The greening of the Northern Great Plains and its biogeochemical precursors. *Glob. Change Biol.* **26**, 5404–5413 (2020). doi: 10.1111/gcb.15115; pmid: 32289875
18. J. Peñuelas, M. Estiarte, Trends in plant carbon concentration and plant demand for N throughout this century. *Oecologia* **109**, 69–73 (1996). doi: 10.1007/s004420050059; pmid: 28307614
19. J. Peñuelas, I. Filella, Herbaria century record of increasing eutrophication in Spanish terrestrial ecosystems. *Glob. Change Biol.* **7**, 427–433 (2001). doi: 10.1046/j.1365-2486.2001.00421.x
20. V. L. DeLeo, D. N. L. Menge, E. M. Hanks, T. E. Juenger, J. R. Lasky, Effects of two centuries of global environmental variation on phenology and physiology of *Arabidopsis thaliana*. *Glob. Change Biol.* **26**, 523–538 (2020). doi: 10.1111/gcb.14880; pmid: 31665819
21. D. Kou *et al*., Progressive nitrogen limitation across the Tibetan alpine permafrost region. *Nat. Commun.* **11**, 3331 (2020). doi: 10.1038/s41467-020-17169-6; pmid: 32620773
22. L. H. Ziska *et al*., Rising atmospheric CO_2 is reducing the protein concentration of a floral pollen source essential for North American bees. *Proc. Biol. Sci.* **283**, 20160414 (2016). doi: 10.1098/rspb.2016.0414; pmid: 27075256
23. J. Penuelas *et al*., Increasing atmospheric CO_2 concentrations correlate with declining nutritional status of European forests. *Commun. Biol.* **3**, 125 (2020). doi: 10.1038/s42003-020-0839-y; pmid: 32170162
24. M. Jonard *et al*., Tree mineral nutrition is deteriorating in Europe. *Glob. Change Biol.* **21**, 418–430 (2015). doi: 10.1111/gcb.12657; pmid: 24920268
25. M. Bauters *et al*., Century-long apparent decrease in intrinsic water-use efficiency with no evidence of progressive nutrient limitation in African tropical forests. *Glob. Change Biol.* **26**, 4449–4461 (2020). doi: 10.1111/gcb.15145; pmid: 32364642
26. P. Hietz *et al*., Long-term change in the nitrogen cycle of tropical forests. *Science* **334**, 664–666 (2011). doi: 10.1126/science.1211979; pmid: 22053047
27. X. Liu *et al*., Enhanced nitrogen deposition over China. *Nature* **494**, 459–462 (2013). doi: 10.1038/nature11917; pmid: 23426264
28. P. M. Groffman *et al*., Nitrogen oligotrophication in northern hardwood forests. *Biogeochemistry* **141**, 523–539 (2018). doi: 10.1007/s10533-018-0445-y
29. K. K. McLauchlan, J. M. Craine, W. W. Oswald, P. R. Leavitt, G. E. Likens, Changes in nitrogen cycling during the past century in a northern hardwood forest. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7466–7470 (2007). doi: 10.1073/pnas.0701779104; pmid: 17446271
30. R. D. Yanai *et al*., From missing source to missing sink: Long-term changes in the nitrogen budget of a northern hardwood forest. *Environ. Sci. Technol.* **47**, 11440–11448 (2013). doi: 10.1021/es4025723; pmid: 24050261
31. K. F. Patel *et al*., Forest N dynamics after 25 years of Whole Watershed N Enrichment: The Bear Brook Watershed in Maine. *Soil Sci. Soc. Am. J.* **83**, 161–174 (2019). doi: 10.2136/sssaj2018.09.0348
32. K. F. Patel, I. J. Fernandez, Nitrogen mineralization in O horizon soils during 27 years of nitrogen enrichment at the Bear Brook Watershed in Maine, USA. *Environ. Monit. Assess.* **190**, 563 (2018). doi: 10.1007/s10661-018-6945-3; pmid: 30167903
33. S. G. McNulty, J. L. Boggs, J. D. Aber, L. E. Rustad, Spruce-fir forest changes during a 30-year nitrogen saturation experiment. *Sci. Total Environ.* **605-606**, 376–390 (2017). doi: 10.1016/j.scitotenv.2017.06.147; pmid: 28668749
34. F. S. Gilliam *et al*., Decreased atmospheric nitrogen deposition in eastern North America: Predicted responses of forest ecosystems. *Environ. Pollut.* **244**, 560–574 (2019). doi: 10.1016/j.envpol.2018.09.135; pmid: 30384062
35. S. Díaz, J. P. Grime, J. Harris, E. McPherson, Evidence of a feedback mechanism limiting plant response to elevated carbon dioxide. *Nature* **364**, 616–617 (1993). doi: 10.1038/364616a0
36. D. R. Taub, X. Wang, Why are nitrogen concentrations in plant tissues lower under elevated CO_2? A critical examination of the hypotheses. *J. Integr. Plant Biol.* **50**, 1365–1374 (2008). doi: 10.1111/j.1744-7909.2008.00754.x; pmid: 19017124
37. C. T. Garten Jr., C. M. Iversen, R. J. Norby, Litterfall ^{15}N abundance indicates declining soil nitrogen availability in a free-air CO_2 enrichment experiment. *Ecology* **92**, 133–139 (2011). doi: 10.1890/10-0293.1; pmid: 21560683
38. R. J. Norby, J. M. Warren, C. M. Iversen, B. E. Medlyn, R. E. McMurtrie, CO_2 enhancement of forest productivity constrained by limited nitrogen availability. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19368–19373 (2010). doi: 10.1073/pnas.1006463107; pmid: 20974944
39. K. M. Becklin, J. S. Medeiros, K. R. Sale, J. K. Ward, Evolutionary history underlies plant physiological responses to global change since the last glacial maximum. *Ecol. Lett.* **17**, 691–699 (2014). doi: 10.1111/ele.12271; pmid: 24636555

40. B. S. Ripley, J. Cunniff, C. P. Osborne, Photosynthetic acclimation and resource use by the C_3 and C_4 subspecies of *Alloteropsis semialata* in low CO_2 atmospheres. *Glob. Change Biol.* **19**, 900–910 (2013). doi: 10.1111/gcb.12091; pmid: 23504846
41. S. P. Long, E. A. Ainsworth, A. Rogers, D. R. Ort, Rising atmospheric carbon dioxide: Plants FACE the future. *Annu. Rev. Plant Biol.* **55**, 591–628 (2004). doi: 10.1146/annurev.arplant.55.031903.141610; pmid: 15377233
42. A. D. B. Leakey *et al.*, Elevated CO_2 effects on plant carbon, nitrogen, and water relations: Six important lessons from FACE. *J. Exp. Bot.* **60**, 2859–2876 (2009). doi: 10.1093/jxb/erp096; pmid: 19401412
43. A. C. Finzi *et al.*, Increases in nitrogen uptake rather than nitrogen-use efficiency support higher rates of temperate forest productivity under elevated CO_2. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14014–14019 (2007). doi: 10.1073/pnas.0706518104; pmid: 17709743
44. Y. Kuzyakov, W. R. Horwath, M. Dorodnikov, E. Blagodatskaya, Review and synthesis of the effects of elevated atmospheric CO_2 on soil processes: No changes in pools, but increased fluxes and accelerated cycles. *Soil Biol. Biochem.* **128**, 66–78 (2019). doi: 10.1016/j.soilbio.2018.10.005
45. Y. Luo *et al.*, Progressive Nitrogen Limitation of Ecosystem Responses to Rising Atmospheric Carbon Dioxide. *Bioscience* **54**, 731 (2004). doi: 10.1641/0006-3568(2004)054[0731:PNLOER]2.0.CO;2
46. M. J. Hovenden *et al.*, Warming prevents the elevated CO_2-induced reduction in available soil nitrogen in a temperate, perennial grassland. *Glob. Change Biol.* **14**, 1018–1024 (2008). doi: 10.1111/j.1365-2486.2008.01558.x
47. W. I. J. Dieleman *et al.*, Simple additive effects are rare: A quantitative review of plant biomass and soil process responses to combined manipulations of CO_2 and temperature. *Glob. Change Biol.* **18**, 2681–2693 (2012). doi: 10.1111/j.1365-2486.2012.02745.x; pmid: 24501048
48. P. B. Reich, J. Oleksyn, Global patterns of plant leaf N and P in relation to temperature and latitude. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11001–11006 (2004). doi: 10.1073/pnas.0403588101; pmid: 15213326
49. P. B. Reich, J. Oleksyn, M. G. Tjoelker, Needle Respiration and Nitrogen Concentration in Scots Pine Populations from a Broad Latitudinal Range: A Common Garden Test with Field-Grown Trees. *Funct. Ecol.* **10**, 768 (1996). doi: 10.2307/2390512
50. M. G. Tjoelker, P. B. Reich, J. Oleksyn, Changes in leaf nitrogen and carbohydrates underlie temperature and CO_2 acclimation of dark respiration in five boreal tree species. *Plant Cell Environ.* **22**, 767–778 (1999). doi: 10.1046/j.1365-3040.1999.00435.x
51. A. J. Elmore, D. M. Nelson, J. M. Craine, Earlier springs are causing reduced nitrogen availability in North American eastern deciduous forests. *Nat. Plants* **2**, 16133 (2016). doi: 10.1038/nplants.2016.133; pmid: 27618399
52. M. A. Dawes, P. Schleppi, S. Hättenschwiler, C. Rixen, F. Hagedorn, Soil warming opens the nitrogen cycle at the alpine treeline. *Glob. Change Biol.* **23**, 421–434 (2017). doi: 10.1111/gcb.13365; pmid: 27207568
53. D. S. Novem Auyeung, V. Suseela, J. S. Dukes, Warming and drought reduce temperature sensitivity of nitrogen transformations. *Glob. Change Biol.* **19**, 662–676 (2013). doi: 10.1111/gcb.12063; pmid: 23504800
54. S. Xu, J. Sardans, J. Zhang, J. Peñuelas, Variations in foliar carbon:nitrogen and nitrogen:phosphorus ratios under global change: a meta-analysis of experimental field studies. *Sci. Rep.* **10**, 12156 (2020). doi: 10.1038/s41598-020-68487-0; pmid: 32699217
55. E. Bai *et al.*, A meta-analysis of experimental warming effects on terrestrial nitrogen pools and dynamics. *New Phytol.* **199**, 441–451 (2013). doi: 10.1111/nph.12252; pmid: 23550663
56. M. J. Hovenden, P. C. D. Newton, K. E. Wills, Seasonal not annual rainfall determines grassland biomass response to carbon dioxide. *Nature* **511**, 583–586 (2014). doi: 10.1038/nature13281; pmid: 24870242
57. L. He *et al.*, Nitrogen Availability Dampens the Positive Impacts of CO_2 Fertilization on Terrestrial Ecosystem Carbon and Water Cycles. *Geophys. Res. Lett.* **44**, 11590–11600 (2017). doi: 10.1002/2017GL075981
58. M. Lee, E. Shevliakova, C. A. Stock, S. Malyshev, P. C. D. Milly, Prominence of the tropics in the recent rise of global nitrogen pollution. *Nat. Commun.* **10**, 1437 (2019). doi: 10.1038/s41467-019-13567-7; pmid: 31796746
59. B. W. Abbott *et al.*, Tundra wildfire triggers sustained lateral nutrient loss in Alaskan Arctic. *Glob. Change Biol.* **27**, 1408–1430 (2021). doi: 10.1111/gcb.15507; pmid: 33394532
60. Z. Zhu *et al.*, Greening of the Earth and its drivers. *Nat. Clim. Change* **6**, 791–795 (2016). doi: 10.1038/nclimate3004
61. T. F. Keenan *et al.*, Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nat. Clim. Change* **4**, 598–604 (2014). doi: 10.1038/nclimate2253
62. B. A. Hungate, J. S. Dukes, M. R. Shaw, Y. Luo, C. B. Field, Nitrogen and climate change. *Science* **302**, 1512–1513 (2003). doi: 10.1126/science.1091390; pmid: 14645831
63. C. Terrer *et al.*, Nitrogen and phosphorus constrain the CO_2 fertilization of global plant biomass. *Nat. Clim. Change* **9**, 684–689 (2019). doi: 10.1038/s41558-019-0545-2
64. P. B. Reich *et al.*, Synergistic effects of four climate change drivers on terrestrial carbon cycling. *Nat. Geosci.* **13**, 787–793 (2020). doi: 10.1038/s41561-020-00657-1
65. S. Wang *et al.*, Recent global decline of CO_2 fertilization effects on vegetation photosynthesis. *Science* **370**, 1295–1300 (2020). doi: 10.1126/science.abb7772; pmid: 33303610
66. T. C. R. White, The importance of a relative shortage of food in animal ecology. *Oecologia* **33**, 71–86 (1978). doi: 10.1007/BF00376997; pmid: 28309267
67. H. L. Throop, M. T. Lerdau, Effects of nitrogen deposition on insect herbivory: Implications for community and ecosystem processes. *Ecosystems* **7**, 109–133 (2004). doi: 10.1007/s10021-003-0225-x
68. F. Slansky Jr., P. Feeny, Stabilization of the Rate of Nitrogen Accumulation by Larvae of the Cabbage Butterfly on Wild and Cultivated Food Plants. *Ecol. Monogr.* **47**, 209–228 (1977). doi: 10.2307/1942617
69. P. Stiling, T. Cornelissen, How does elevated carbon dioxide (CO_2) affect plant-herbivore interactions? A field experiment and meta-analysis of CO_2-mediated changes on plant chemistry and herbivore performance. *Glob. Change Biol.* **13**, 1823–1842 (2007). doi: 10.1111/j.1365-2486.2007.01392.x
70. M. J. Scriber, P. Feeny, Growth of Herbivorous Caterpillars in Relation to Feeding Specialization and to the Growth Form of Their Food. *Ecology* **60**, 829–850 (1979). doi: 10.2307/1936618
71. T. Ylioja, H. Roininen, M. P. Ayres, M. Rousi, P. W. Price, Host-driven population dynamics in an herbivorous insect. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10735–10740 (1999). doi: 10.1073/pnas.96.19.10735; pmid: 10485895
72. C. S. Awmack, S. R. Leather, Host plant quality and fecundity in herbivorous insects. *Annu. Rev. Entomol.* **47**, 817–844 (2002). doi: 10.1146/annurev.ento.47.091201.145300; pmid: 11729092
73. J. Pöyry *et al.*, The effects of soil eutrophication propagate to higher trophic levels. *Glob. Ecol. Biogeogr.* **26**, 18–30 (2017). doi: 10.1111/geb.12521
74. J. Cebrian *et al.*, Producer nutritional quality controls ecosystem trophic structure. *PLOS ONE* **4**, e4929 (2009). doi: 10.1371/journal.pone.0004929; pmid: 19300514
75. E. A. R. Welti, K. A. Roeder, K. M. de Beurs, A. Joern, M. Kaspari, Nutrient dilution and climate cycles underlie declines in a dominant insect herbivore. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 7271–7275 (2020). doi: 10.1073/pnas.1920012117; pmid: 32152101
76. R. van Klink *et al.*, Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science* **368**, 417–420 (2020). doi: 10.1126/science.aax9931; pmid: 32327596
77. C. A. Deans *et al.*, Nutrition affects insect susceptibility to Bt toxins. *Sci. Rep.* **7**, 39705 (2017). doi: 10.1038/srep39705; pmid: 28045087
78. M. P. Nessel, T. Konnovitch, G. Q. Romero, A. L. González, Nitrogen and phosphorus enrichment cause declines in invertebrate populations: A global meta-analysis. *Biol. Rev.* **96**, 2617–2637 (2021). doi: 10.1111/brv.12771; pmid: 34173704
79. S. W. Kim *et al.*, Meeting Global Feed Protein Demand: Challenge, Opportunity, and Strategy. *Annu. Rev. Anim. Biosci.* **7**, 221–243 (2019). doi: 10.1146/annurev-animal-030117-014838; pmid: 30418803
80. J. M. Craine, A. Elmore, J. P. Angerer, Long-term declines in dietary nutritional quality for North American cattle. *Environ. Res. Lett.* **12**, 044019 (2017). doi: 10.1088/1748-9326/aa67a4
81. J. D. Derner, L. Hunt, E. Filho, J. Ritten, J. Capper, G. Han, in *Rangeland Systems* (Springer Nature, 2017), pp. 347–372.
82. J. M. Craine, E. G. Towne, M. Miller, N. Fierer, Climatic warming and the future of bison as grazers. *Sci. Rep.* **5**, 16738 (2015). doi: 10.1038/srep16738; pmid: 26567987
83. W. R. Hill, J. Rinchard, S. Czesny, Light, nutrients and the fatty acid composition of stream periphyton. *Freshw. Biol.* **56**, 1825–1836 (2011). doi: 10.1111/j.1365-2427.2011.02622.x
84. M. Torres-Ruiz, J. D. Wehr, A. A. Perrone, Trophic relations in a stream food web: Importance of fatty acids for macroinvertebrate consumers. *J. N. Am. Benthol. Soc.* **26**, 509–522 (2007). doi: 10.1899/06-070.1
85. R. D. Sabo *et al.*, Positive correlation between wood $\delta^{15}N$ and stream nitrate concentrations in two temperate deciduous forests. *Environ. Res. Commun.* **2**, 025003 (2020). doi: 10.1088/2515-7620/ab77f8
86. H. A. de Wit *et al.*, Land-use dominates climate controls on nitrogen and phosphorus export from managed and natural Nordic headwater catchments. *Hydrol. Processes* **34**, 4831–4850 (2020). doi: 10.1002/hyp.13939
87. F. Sánchez-Bayo, K. A. G. Wyckhuys, Further evidence for a global decline of the entomofauna. *Austral Entomol.* **60**, 9–26 (2021). doi: 10.1111/aen.12509
88. C. M. Duarte, D. J. Conley, J. Carstensen, M. Sánchez-Camacho, Return to Neverland: Shifting baselines affect eutrophication restoration targets. *Estuaries Coasts* **32**, 29–36 (2009). doi: 10.1007/s12237-008-9111-2
89. A. Oczkowski *et al.*, How the distribution of anthropogenic nitrogen has changed in Narragansett Bay (RI, USA) following major reductions in nutrient loads. *Estuaries Coasts* **41**, 2260–2276 (2018). doi: 10.1007/s12237-018-0435-2; pmid: 30971866
90. S. W. Nixon, Replacing the Nile: Are anthropogenic nutrients providing the fertility once brought to the Mediterranean by a great river? *Ambio* **32**, 30–39 (2003). doi: 10.1579/0044-7447-32.1.30; pmid: 12691489
91. C. Oviatt *et al.*, Managed nutrient reduction impacts on nutrient concentrations, water clarity, primary production, and hypoxia in a north temperate estuary. *Estuar. Coast. Shelf Sci.* **199**, 25–34 (2017). doi: 10.1016/j.ecss.2017.09.026
92. C. L. Meunier *et al.*, From elements to function: Toward unifying ecological stoichiometry and trait-based ecology. *Front. Environ. Sci.* **5**, 18 (2017). doi: 10.3389/fenvs.2017.00018
93. J. Meyerholt, K. Sickel, S. Zaehle, Ensemble projections elucidate effects of uncertainty in terrestrial nitrogen limitation on future carbon uptake. *Glob. Change Biol.* **26**, 3978–3996 (2020). doi: 10.1111/gcb.15114; pmid: 32285534
94. E. Pennisi, Carbon dioxide increase may promote 'insect apocalypse'. *Science* **368**, 459 (2020). doi: 10.1126/science.368.6490.459; pmid: 32355011
95. M. L. Word *et al.*, Soil-targeted interventions could alleviate locust and grasshopper pest pressure in West Africa. *Sci. Total Environ.* **663**, 632–643 (2019). doi: 10.1016/j.scitotenv.2019.01.313; pmid: 30731409
96. J. E. Compton *et al.*, Ecological and water quality consequences of nutrient addition for salmon restoration in the Pacific Northwest. *Front. Ecol. Environ.* **4**, 18–26 (2006). doi: 10.1890/1540-9295(2006)004[0018:EAWQCO]2.0.CO;2
97. J. Sardans *et al.*, Changes in nutrient concentrations of leaves and roots in response to global change factors. *Glob. Change Biol.* **23**, 3849–3856 (2017). doi: 10.1111/gcb.13721; pmid: 28407324
98. L. Deng *et al.*, Soil GHG fluxes are altered by N deposition: New data indicate lower N stimulation of the N_2O flux and greater stimulation of the calculated C pools. *Glob. Change Biol.* **26**, 2613–2629 (2020). doi: 10.1111/gcb.14970
99. M. Engardt, D. Simpson, M. Schwikowski, L. Granat, Deposition of sulphur and nitrogen in Europe 1900–2050. Model calculations and comparison to historical observations. *Tellus B* **69**, 1328945 (2017). doi: 10.1080/16000889.2017.1328945
100. S. Running, S. Zhao, MOD17A3HGF MODIS/Terra Net Primary Production Gap-Filled Yearly L4 Global 500 m SIN Grid V006. NASA EOSDIS Land Processes DAAC (2019). doi: 10.5067/MODIS/MOD17A3HGF.006

ACKNOWLEDGMENTS

We are grateful to A. Carew for graphic design assistance and to J. Penuelas, R. Sabo, M. Engardt, and D. Simpson for providing data. This paper also benefitted from discussions with J. Brookshire, J. Campbell, C. Cleveland, E. Davidson, J. Dukes, M. Green, B. Hungate, Y. Luo, B. Poulter, B. Shuman, C. Terrer, and E. Welti. We appreciate the thoughtful comments of three anonymous reviewers, whose perspectives improved this Review. **Funding:** This work was supported by the National Socio-Environmental Synthesis Center (SESYNC) under funding received from the National Science Foundation (NSF) DBI-1639145 (R.E.M., Q.D.R., and A.J.E.). Long-term nitrogen studies at the Hubbard Brook Experimental Forest have been supported by the NSF Long-Term Ecological Research program since 1988 (P.M.G.). Work on this paper was also supported, in part, by the US Department of Agriculture, Agricultural Research Service (J.A.); US Department of Agriculture, Forest Service, Northern Research Station (N.K.L.); NSF grants 1638688, 1920908, and 832210 (S.V.O.); and RI Sea Grant (R.W.F.). **Competing interests:** The authors declare that they have no competing interests

10.1126/science.abh3767

Science
JOURNALS AAAS

Publish your research in the *Science* family of journals

The *Science* family of journals (*Science, Science Advances, Science Immunology, Science Robotics, Science Signaling*, and *Science Translational Medicine)* are among the most highly-regarded journals in the world for quality and selectivity. Our peer-reviewed journals are committed to publishing cutting-edge research, incisive scientific commentary, and insights on what's important to the scientific world at the highest standards.

Submit your research today!

Learn more at **Science.org/journals**

RESEARCH ARTICLE SUMMARY

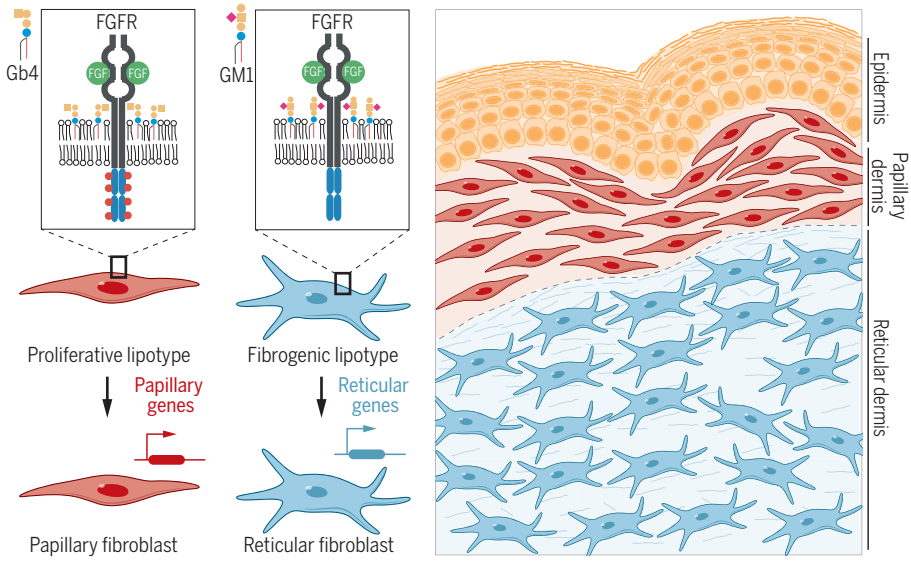
LIPIDOMICS

Sphingolipids control dermal fibroblast heterogeneity

Laura Capolupo, Irina Khven, Alex R. Lederer, Luigi Mazzeo, Galina Glousker, Sylvia Ho, Francesco Russo, Jonathan Paz Montoya, Dhaka R. Bhandari, Andrew P. Bowman, Shane R. Ellis, Romain Guiet, Olivier Burri, Johanna Detzner, Johannes Muthing, Krisztian Homicsko, François Kuonen, Michel Gilliet, Bernhard Spengler, Ron M. A. Heeren, G. Paolo Dotto, Gioele La Manno*, Giovanni D'Angelo*

INTRODUCTION: External signals (e.g., hormones, cytokines, and growth factors) and cell-autonomous properties (e.g., the transcriptional and metabolic states of individual cells) concur to determine cell-fate decisions. Although the mode of action of external signals has been detailed extensively in decades of intense research, the molecular bases of cell-autonomous contribution to cell-fate decisions have been traditionally more elusive. Lipids are fundamental constituents of all living beings. They participate in energy metabolism, account for the assembly of biological membranes, act as signaling molecules, and interact with proteins to influence their function and intracellular distribution. Eukaryotic cells produce thousands of different lipids, each endowed with peculiar structural features and contributing to specific biological functions. With the development of lipidomics, we can now understand the lipid compositional complexity of cells and start making sense of lipidome dynamics. Lipidomes indeed vary among cell types and are reprogrammed in differentiation events. However, whether and how lipidome remodeling assists changes in cell identity is not understood.

RATIONALE: Human dermal fibroblasts are cell constituents of our skin that display cell-to-cell phenotypic heterogeneity as a result of their dynamic cell identity. Thus, individual dermal fibroblasts can adopt different cell specializations that are responsible for wound repair, fibrosis, or remodeling of the extracellular matrix. Whether lipid metabolism is differently shaped in fibroblasts with different phenotypes and if lipid composition participates in the establishment of fibroblast subtypes were unknown. Here, we addressed both the overall lipid composition and phenotypic states of hundreds of individual dermal fibroblasts looking for a possible role of lipids in the determination of dermal fibroblast identity.



Sphingolipids control dermal fibroblast heterogeneity. Human dermal fibroblasts exist in multiple lipid configurations (lipotypes) marked by different sphingolipids. Sphingolipids such as Gb4 or GM1, distinctly modulate FGF receptor (FGFR) signaling upon exposure to FGF2. As a result of this modulation, lipotypes promote alternative transcriptional programs that are associated with papillary or reticular fibroblasts. Accordingly, fibroblasts bearing different lipotypes populate the reticular and papillary layers of the skin.

RESULTS: We coupled high-resolution mass spectrometry imaging and single-cell mRNA sequencing to resolve both lipidomes and transcriptomes of individual dermal fibroblasts. We found that dermal fibroblasts exist in multiple lipid compositional states that correspond to transcriptional subpopulations in vitro and to fibroblasts populating different layers of the skin in vivo. We isolated the metabolic pathways that account for this correlation and found that sphingolipids are major markers of the different lipid compositional states that we named lipotypes. We also found that lipotype heterogeneity influences cell identity by diversifying the response of otherwise identical cells to extracellular stimuli and that manipulating sphingolipid composition is sufficient to reprogram cells toward different phenotypic states. We also found that lipid composition and signaling pathways are wired in self-sustained circuits that account for the metabolic and transcriptional fibroblast heterogeneity. Specifically, we observed that sphingolipids modulate fibroblast growth factor 2 (FGF2) signaling, with globo-series sphingolipids acting as positive regulators and ganglio-series glycosphingolipids as negative regulators. In turn, FGF2 signaling counteracts ganglioside production by sustaining the alternative metabolic pathway leading to the production of globo-series sphingolipids.

CONCLUSION: By studying the lipid composition of individual cells, we found that lipids play a driving role in the determination of cell states. We indeed uncovered an unexpected relationship between lipidomes and transcriptomes in individual cells. In fact, our results indicate that the acquisition of specific lipotypes influenced the activity of signaling receptors and fostered alternative transcriptional states. Cell states are intermediates in the process of cell differentiation in which state switches precede terminal commitment. As a consequence, lipidome remodeling could work as an early driver in the establishment of cell identity, and following lipid metabolic trajectories of individual cells could have the potential to inform us about key mechanisms of cell fate decision. Thus, this study stimulates new questions about the role of lipids in cell-fate decisions and adds a new regulatory component to the self-organization of multicellular systems. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: gioele.lamanno@epfl.ch (G.L.M.); giovanni.dangelo@epfl.ch (G.D.)
Cite this article as L. Capolupo *et al.*, *Science* **376**, eabh1623 (2022). DOI: 10.1126/science.abh1623

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abh1623

RESEARCH ARTICLE

LIPIDOMICS

Sphingolipids control dermal fibroblast heterogeneity

Laura Capolupo[1], Irina Khven[2], Alex R. Lederer[2], Luigi Mazzeo[3], Galina Glousker[4], Sylvia Ho[1], Francesco Russo[5], Jonathan Paz Montoya[1], Dhaka R. Bhandari[6], Andrew P. Bowman[7], Shane R. Ellis[7,8,9], Romain Guiet[10], Olivier Burri[10], Johanna Detzner[11], Johannes Muthing[11], Krisztian Homicsko[12,13,14], François Kuonen[15], Michel Gilliet[15], Bernhard Spengler[6], Ron M. A. Heeren[7], G. Paolo Dotto[16,17,18], Gioele La Manno[2]*, Giovanni D'Angelo[1,5]*

Human cells produce thousands of lipids that change during cell differentiation and can vary across individual cells of the same type. However, we are only starting to characterize the function of these cell-to-cell differences in lipid composition. Here, we measured the lipidomes and transcriptomes of individual human dermal fibroblasts by coupling high-resolution mass spectrometry imaging with single-cell transcriptomics. We found that the cell-to-cell variations of specific lipid metabolic pathways contribute to the establishment of cell states involved in the organization of skin architecture. Sphingolipid composition is shown to define fibroblast subpopulations, with sphingolipid metabolic rewiring driving cell-state transitions. Therefore, cell-to-cell lipid heterogeneity affects the determination of cell states, adding a new regulatory component to the self-organization of multicellular systems.

The division of labor is a fundamental organizational principle of multicellular organisms that is implemented through transcriptional programs resulting in cell types. However, phenotypic heterogeneity can occur across cells of the same type, resulting in different cell states (*1*–*3*). These varying cell states can have physiological significance such as priming diverging differentiation programs (*4*) or contributing to distinct cellular tasks in physiological processes (*5*).

Fibroblasts are a cell type that can plastically transition across multiple states (*6*–*13*). Changes in the proportion of fibroblast subpopulations are associated with fibrosis and contribute to a tissue microenvironment permissive for cancer growth (*14*–*18*). Cell lineage, soluble factors, and the microenvironment (*6*) all contribute to the determination of fibroblast states (*15*), yet the molecular circuits that govern this fibroblast heterogeneity and plasticity have not been fully clarified.

Metabolic rewiring is inherent to cell-fate transitions (*19*), and several metabolic switches involving lipids are important for multicellular organism development (*20*). Nonetheless, only a few studies have investigated lipid composition at the single-cell level and the relevance of its variability (*21*–*24*). Thus, whether lipid metabolism has a role in the determination of cell states remains unclear. Specifically, although lipids modulate the differentiation of stem cells in the skin (*25*), whether and how lipid metabolism participates in fibroblast state plasticity has not been addressed.

Mass spectrometry (MS) techniques now have enough sensitivity to enable single-cell lipidomics (*26*–*28*). In particular, matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) provides coverage of the lipid mass-to-charge-number (m/z) range, causes minimal fragmentation, and has reached a spatial resolution compatible with single-cell analysis while maintaining mass resolution and accuracy (*29*–*36*).

MALDI-MSI reveals the organizing principles of lipid heterogeneity

We performed space-resolved (25 to 50 μm^2 pixel size) MALDI-MSI on low-passage primary dermal human fibroblasts (dHFs) (Fig. 1A). Lipid images (Fig. 1B) were extracted from raw data and lipid identity was attributed (*37*) and validated by electrospray ionization liquid chromatography–mass spectrometry (ESI-LC/MS) (*37*) and multiple reaction monitoring (MRM)–based lipidomics (Fig. 1A; fig S1, A and B; and table S1). Specific attributions were disambiguated by comparison with pure standards (fig. S1C) and targeted LC-MS/MS (fig. S1D). Overall, images of 205 annotated lipids were obtained (*37*) (table S1), which account for a sizable fraction of the dHF lipidome as detected by LC-MS.

The intensities of all the m/z peaks at each scanned location (i.e., pixel) were used to perform a multivariate analysis. Principal component analysis (PCA) revealed that 95% of the pixel-to-pixel variability could be explained by eight principal components (PCs) (Fig. 1C and fig. S1E). The in situ visualization of the PC coordinates corresponding to each pixel delineated distinct distribution patterns for different groups of lipids (Fig. 1C).

PC1 coordinates changed from the inner part of the cell toward the cell periphery, suggesting that this axis captures fundamental differences in lipid composition of the perinuclear and peripheral cell membranes (Fig. 1D). In contrast to what was observed for PC1, PC2 to PC8 coordinates distributed differently among cells, with some cells displaying exclusively positive or negative pixels (Fig. 1C). Lipids belonging to the sphingolipid pathway [i.e., ceramides (Cers), sphingomyelins (SMs), hexosylceramides (HexCers), trihexosylceramides (Gb3s), and globosides (Gb4s)] accounted for these axes of cell-to-cell variation (Fig. 1D and fig. S1E). This confirms previous observations concerning the cell-to-cell variability of specific sphingolipids (*21*, *24*) and extends them to most of the lipid species observed in this pathway. From these results, we conclude that two coexisting axes of lipid variation exist in dHFs. One axis pertains to intracellular organization (*38*) and the other to lipid-related intercellular heterogeneity (*39*).

Single-cell analysis reveals lipid coregulation

To understand the nature of this lipid intercellular heterogeneity, we used optical images to guide cell segmentation and transferred them onto the MS images to obtain a total of 257 single-cell lipidomes from three independent MALDI-MSI recordings (Fig. 2A). After data normalization and batch correction, the cell-to-cell variability associated with individual lipid species was summarized by computing their coefficient of variation (CV) (*37*) across the cell population. The obtained values were used to rank lipids according to their

[1]Interfaculty Institute of Bioengineering and Global Health Institute, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. [2]Brain Mind Institute, Faculty of Life Sciences, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. [3]Department of Biochemistry, University of Lausanne, CH-1066 Epalinges, Switzerland. [4]School of Life Sciences, Swiss Institute for Experimental Cancer Research, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. [5]Institute of Biochemistry and Cellular Biology, National Research Council of Italy, 80131 Napoli, Italy. [6]Institute for Inorganic and Analytical Chemistry, Justus Liebig University Giessen, 35392 Giessen, Germany. [7]Maastricht MultiModal Molecular Imaging Institute, Division of Imaging Mass Spectrometry, Maastricht University, 6629 ER Maastricht, Netherlands. [8]Molecular Horizons and School of Chemistry and Molecular Bioscience, University of Wollongong, Wollongong, New South Wales 2522, Australia. [9]Illawarra Health and Medical Research Institute, Wollongong, New South Wales 2522, Australia. [10]Faculté des Sciences de la Vie, Bioimaging and Optics Platform, École Polytechnique Fédérale de Lausanne, Lausanne, CH-1015 Vaud, Switzerland. [11]Institute of Hygiene, University of Münster, D-48149 Münster, Germany. [12]Department of Oncology, Centre Hospitalier Universitaire Vaudois, CH-1011 Lausanne, Switzerland. [13]Swiss Cancer Center Leman, CH-1015 Lausanne, Switzerland. [14]The Ludwig Institute for Cancer Research, Lausanne Branch, CH-1066 Epalinges, Switzerland. [15]Département de Dermatologie et Vénéréologie, Centre Hospitalier Universitaire Vaudois, CH-1011 Lausanne, Switzerland. [16]Personalized Cancer Prevention Research Unit, Head and Neck Surgery Division, Centre Hospitalier Universitaire Vaudois, CH-1011 Lausanne, Switzerland. [17]Department of Biochemistry, University of Lausanne, CH-1066 Epalinges, Switzerland. [18]Cutaneous Biology Research Center, Massachusetts General Hospital, Charlestown, MA 02129, USA.
*Corresponding author. Email: gioele.lamanno@epfl.ch (G.L.M.); giovanni.dangelo@epfl.ch (G.D.)

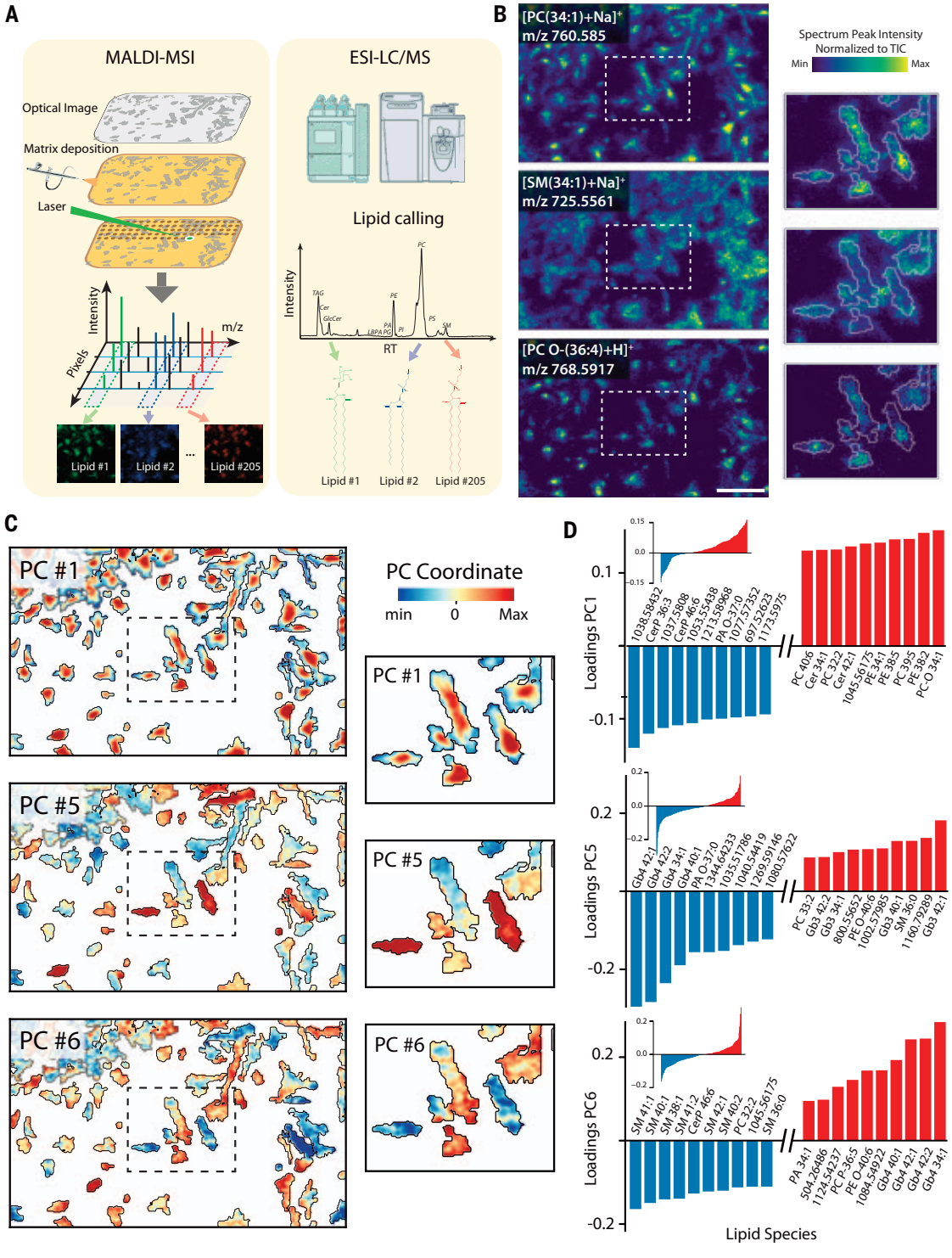


Fig. 1. Single-pixel MALDI-MSI analysis on dHFs. (**A**) Schematic drawing of MALDI-MSI workflow. Cells were fixed, matrix was deposited, and MALDI-MSI was performed by rasterizing the laser across a selected area. For each spot, a mass spectrum was collected and mass images were obtained for each ion by plotting *m/z* intensity at the corresponding *x* and *y* coordinates (left panel). For peak identification, total lipid extracts were analyzed by ESI-LC/MS (right panel). Lipids identified by ESI-LC/MS were then compared with the ones obtained by MALDI-MSI. (**B**) Ion images (50 μm^2/pixel; 354 × 218 pixels) of selected lipids recorded in positive-ion mode. Insets show individual cells images at higher magnification. $[PC(34:1)+Na]^+$, phosphatidylcholine with acyl chains consisting of 34 carbon atoms and one double bond complexed with sodium; $[SM(34:1)+Na]^+$, SM with a backbone of 34 carbon atoms and one double bond complexed with sodium; $[PC\ O\text{-}(36:4)+H]^+$, phosphatidylcholine plasmalogen with acyl chains consisting of 36 carbon atoms and four double bonds complexed with hydrogen; TIC, total ion current. Scale bar, 500 µm. (**C**) Images displaying at each location the PCA coordinate of each pixel. PC1, PC5, and PC6 values are displayed using a divergent color map; positive coordinates are shown in red and negative in blue. Insets show individual cell images at higher magnification. (**D**) Bar plots showing the contribution of the top 10 lipids with higher (red) and lower (blue) loadings for each PC. Miniatures in the upper left corner show the entire distribution.

decreasing CV. Sphingolipids populated the top CV ranking positions, confirming that the sphingolipid pathway (fig. S1F) is subjected to high cell-to-cell variability (Fig. 2B).

To test whether the sphingolipid metabolic pathway is coordinately modulated or if different sphingolipid subsets are controlled independently in different cells, we created a pairwise lipid-lipid correlation (Pearson's *R*) matrix (Fig. 2C). Although phospholipid species did not form biochemically meaningful cliques, sphingolipids were clustered in groups consisting of compounds bearing the same head group (i.e., OH with Cers, hexose with HexCers, trihexose with Gb3s, and *N*-acetyl-hexose-trihexose with Gb4s) but with different Cer backbones (mostly 34:1, 40:1, 42:1, and 42:2) (fig. S1F). This suggests that specific enzymatic

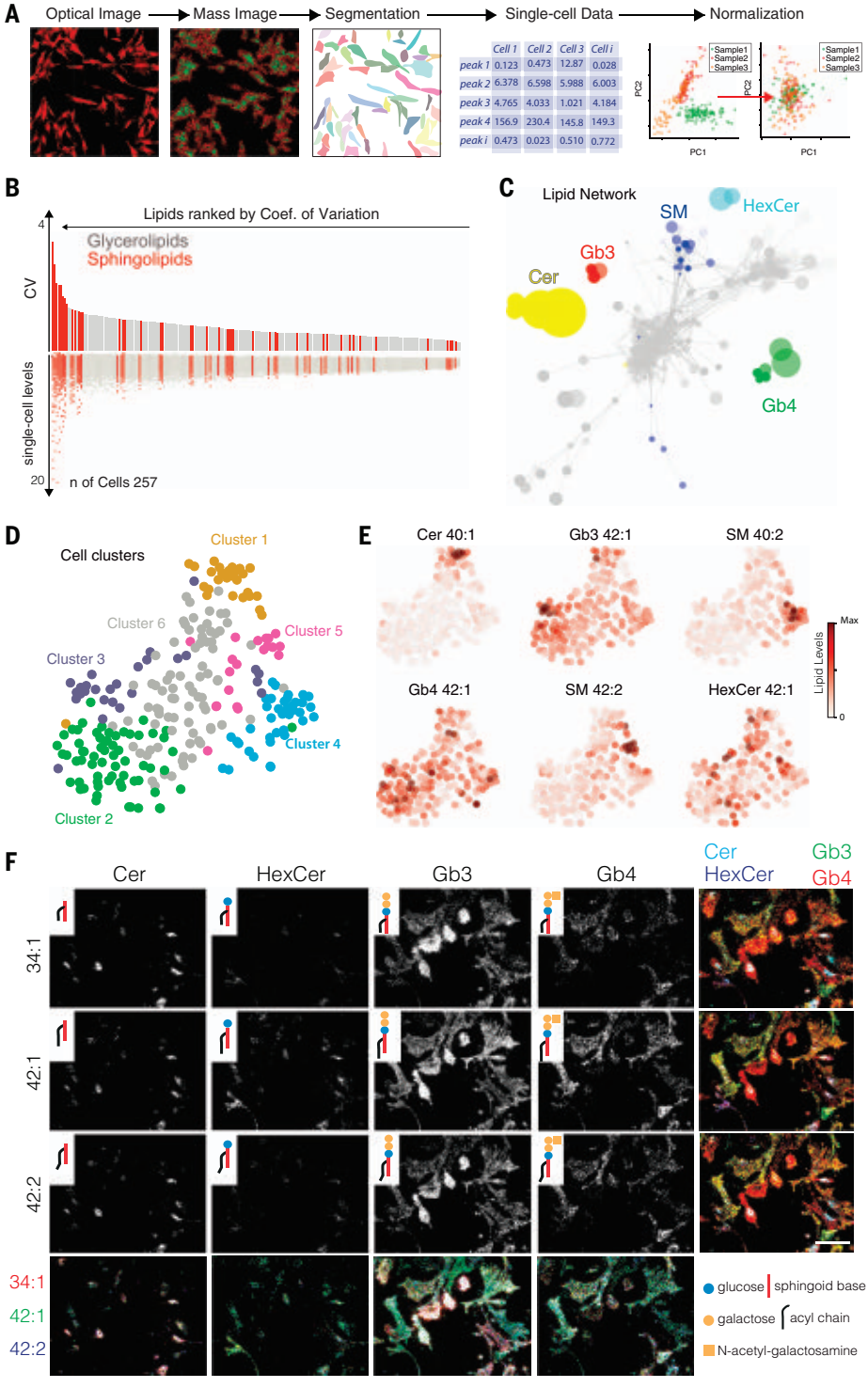


Fig. 2. Single-cell lipidomics analysis. (**A**) Schematic of the approach used for single-cell analysis of MALDI-MSI data. Confocal micrographs were used as guides to segment cells out of the mass images, and single-cell ion abundance was computed as the TIC-normalized peak intensity. Different acquisitions were combined after ComBat batch correction. (**B**) Barplot showing the CV of lipids computed across 257 cells. Lipids were ranked by CV and color coded according to their class (sphingolipids are shown in red, glycerolipids in gray). Single-cell lipid levels are shown in the bottom part of the plot. (**C**) Lipid covariation network. Nodes represent individual lipids, size is proportional to the CV, and color is according to lipid class: Cers are shown in yellow, Gb3s in red, SMs in blue, HexCers in cyan, and Gb4s in green. Edges connect two lipids where the correlation coefficient is >0.85. (**D**) t-distributed stochastic neighbor embedding (t-SNE) of the single-cell lipidomics data. Cells are colored by the clusters defined by hierarchical clustering. (**E**) t-SNE colored by the abundance of sphingolipids. (**F**) Mass images showing the spatial distribution of sphingolipid precursors (Cers and HexCers) and complex sphingolipids (Gb3s and Gb4s) composed of different backbones (34:1, 42:1, and 42:2). Miniatures in the top left corner of each image depict a simplified schematic of the lipid structure (compare with fig. S2A). Scale bar, 500 μm.

activities responsible for Cer processing, rather than those producing different Cer backbones, are variable at a cell-to-cell level. Accordingly, the relative abundances of lipids sharing the same head group were more correlated than those sharing the same backbone (fig. S1G).

Single-cell lipidomes were then used to group cells according to their lipid composition (*37*), resulting in distinct cell clusters (Fig. 2D and fig. S1H). When the levels of sphingolipids were considered, we observed that certain species (i.e., Cers, HexCers, Gb3s, and Gb4s) were enriched in specific cell clusters, suggesting that dHFs exist in distinct sphingolipid metabolic states (Fig. 2, E and F).

Sphingolipids define dHF lipotypes

To validate these results, we stained cells with fluorescently labeled bacterial toxins that recognize different sphingolipid head groups: Shiga toxin 1a (ShTxB1a) binds to Gb3 (*40*), Shiga toxin 2e (ShTxB2e) binds to Gb3 and Gb4 (*41*), and Cholera toxin B (ChTxB) binds the ganglioside GM1 (*42*). Toxins stained dHFs with a pattern reminiscent of the variability observed by MALDI-MSI (Fig. 3A and fig. S2A). Treatment with inhibitors of sphingolipid production [fumonisin B1 (FB1) (*43*) and D-threo-1-phenyl-2-decanoylamino-3-morpholino-1-propanol (D-PDMP) (*44*)] or silencing the expression of

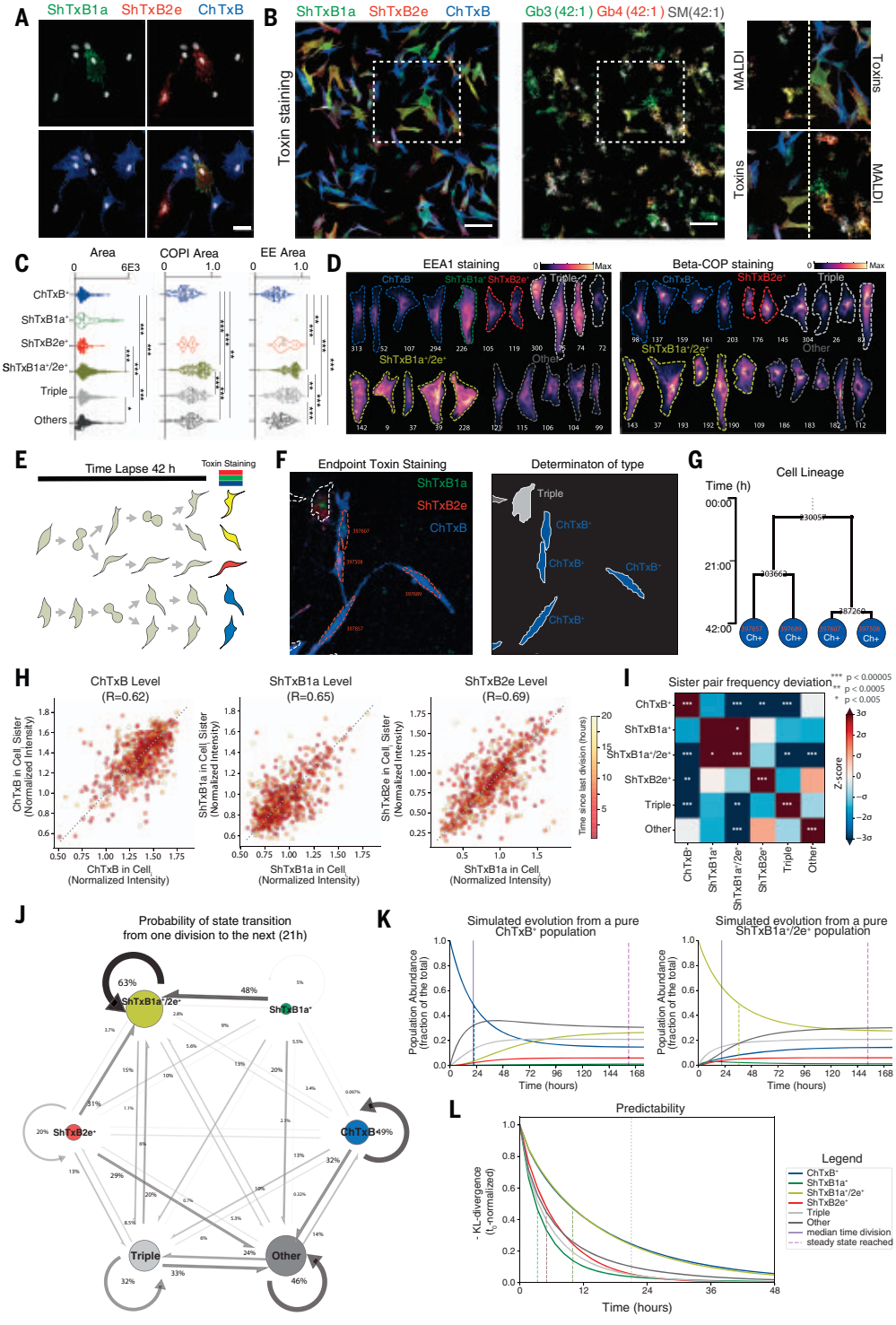


Fig. 3. Identification of dHF lipotypes by MALDI-MSI and toxin staining. (**A**) Confocal micrographs showing cells stained with bacterial toxins ShTxB1a (green), ShTxB2e (red), ChTxB (blue), and Hoechst (gray) for nuclei. Scale bar, 50 μm. (**B**) Side-by-side comparison of toxin staining and MALDI-MSI acquisition on the same cells. First, cells were stained with bacterial toxins as in (A), and images were acquired by confocal microscopy (left panel). Then, MALDI-MSI (25 μm^2/pixel) was performed on the same cells (center panel). Mass images (320 × 320 pixels) of complex sphingolipids [SM (42:1), Gb3 (42:1), and Gb4 (42:1)] are shown. Scale bar, 200 μm. (**C**) Cells were stained with bacterial toxins as in (A) and with antibodies against Beta-COP (COPI vesicles) or EEA1 [early endosomes (EEs)], and images were acquired by confocal microscopy. Normalized fluorescence intensities of toxin and organelle marker stainings of single cells were used to analyze the correlation between lipotypes and cell area and with exo/endocytic organelles. Data are shown as violin plots. *P < 0.05, **P < 0.01, ***P < 0.001, ordinary one-way ANOVA. (**D**) Representative cells stained as for EEA1 or Beta-COP and classified according to their lipotypes. (**E**) Schematic representation of dHF cell lineage tracking. (**F**) Representative confocal micrograph of toxin-stained dHFs before (left) and after (right) segmentation with Cellpose. Segmented cell colors correspond to the different lipotypes. (**G**) Lineage reconstruction for the cells illustrated in (F) as inferred using TrackMate (*37*). (**H**) Correlation plots of normalized ChTxB, ShTxB1a, and ShTxB2e intensities between daughter cells at the time course end point. Dots are colored by the number of hours after mother cell division. (**I**) Heatmap of frequencies for two lipotypes occurring in two sister cells colored by z-score. Positive deviation from zero indicates an increased observed frequency of the sister-lipotype combination compared with random chance, and negative deviation from zero indicates a decreased observed frequency of the sister-lipotype combination compared with random chance. P values were calculated using the bootstrap pairwise t test (*37*). (**J**) Probability of a lipotype state transition occurring in a cell over a 21-hour time period as estimated using CELLMA (*37*). Probabilities are located at the corresponding arrow tails. (**K**) Markov model–simulated evolution of a pure ChTxB$^+$ (left) or pure ShTxB1a$^+$/2e$^+$ (right) cell population over 7 days. (**L**) Line plots displaying the evolution of the state predictability of a cell (or its progeny) after a certain time from an original state measurement. Differently colored tracks correspond to a different original lipotype measurement (t = 0). Kullback-Leibler divergence is evaluated between the probability distribution vector obtained using the Markov transition matrix and the steady-state probability distribution (i.e., the best uninformed guess).

B4GALT5 encoding lactosylceramide synthase (LCS) reduced toxin binding (fig. S2, B and C) without inducing significant toxicity (fig. S2D), indicating that toxins are a faithful proxy for dHF sphingolipid composition in our setting.

As further validation, dHFs were first fixed and stained with toxins and then imaged by MALDI-MSI (Fig. 3B). ShTxB1a staining correlated best with Gb3 levels, and ShTxB2e staining correlated well with Gb3 and Gb4 levels, whereas neither correlated with SM levels. ChTxB staining, our proxy for the levels of GM1 (*42*), a sphingolipid not detected by MALDI-MSI in positive-ion mode, did not correlate with any of the sphingolipids detected

by mass imaging (fig. S2, E and F). We then analyzed dHFs derived from four unrelated healthy individuals that displayed an analogous pattern of cell-to-cell sphingolipid variability (fig. S2G), suggesting that bacterial toxins capture cell-to-cell sphingolipid heterogeneity and that sphingolipid heterogeneity is common to dHFs from different individuals.

We categorized dHFs depending on their sphingolipid configurations into ChTxB$^+$, ShTxB1a$^+$, ShTxB2e$^+$, ShTxB1a$^+$/2e$^+$, triple$^+$, and "other" (accounting for all other configurations) (fig. S3A). When looking at features associated with these categories (Fig. 3C and fig. S3B), we observed that ShTxB1a$^+$/2e$^+$ and triple$^+$ cells were larger than ChTxB$^+$ and ShTxB2e$^+$ cells and that ShTxB1a$^+$/2e$^+$ had a more complex shape than ChTxB$^+$ cells (Fig. 3C and fig. S3B). We also considered the cell-to-cell variability associated with exo/endocytic organelles (*45*), where sphingolipid production and turnover take place (*46*). We observed that ShTxB1a$^+$/2e$^+$ dHFs have an expanded early endosomal compartment compared with other configurations, with ChTxB$^+$ dHFs showing an opposite phenotype. Similar, although less striking, changes were observed when looking at coat protein complex I (COPI) vesicles and at the Golgi complex (Fig. 3, C and, D, and fig. S3, C and D). Thus, dHFs exist in different lipid metabolic configurations that correspond to distinct cell phenotypes involving cell size and shape and are endowed with different endocytic and secretory states.

To assess the dynamics of sphingolipid configurations, dHF lineages were followed by live microscopy, and individual cells were analyzed by toxin staining after fixation (Fig. 3, E to G; fig. S4A; and movie S1). When the intensity levels associated with individual toxins were considered in pairs of sister cells, we noticed that they were correlated (Fig. 3H) and that lineage-related cells had a higher probability of sharing the same sphingolipid configuration than would be expected by chance (Fig. 3I and fig. S4, B to D). Considering the toxin-staining patterns of lineage-related cells, we modeled the dynamics of lipid configuration switches by developing the cell-state transition estimation by lineage leaf-state Markov analysis (CELLMA) algorithm (*37*). This model predicted that ShTxB1a$^+$/2e$^+$, ChTxB$^+$, and triple$^+$ are stable states with a 37%, 51%, and 68% probability of converting into a different lipid configuration during a single-cell replication cycle (21 hours), respectively. Conversely, ShtxB1a$^+$ and ShtxB2e$^+$ states were more transient and showed a greater propensity (95 and 80% during a single replication, respectively) for converting into ShTxB1a$^+$/2e$^+$ or into "other" lipid configurations (Fig. 3J and fig. S4E).

These dynamics translate into lipid-state transition fluxes such that the predominant lipid configurations are propagated across cell generations (fig. S4F). Accordingly, our model predicts that populations composed of cells all belonging to the same lipid category would revert slowly (i.e., within 7 days) to a heterogeneous steady state (Fig. 3K and fig. S4G). In agreement with this prediction, when we selected ShTxB1a$^+$ or ChTxB$^+$ cells by fluorescence-activated cell sorting (FACS) and kept them in culture for 10 days, the cell cultures reverted to heterogeneous cell populations with lipid-state compositions similar to those from which they were originally selected (fig. S4H).

On the basis of these results, we conclude that dHFs exist in metastable sphingolipid metabolic configurations (Fig. 3L) that correspond to given phenotypic states and persist during cell generations. Hereafter, we refer to these lipid metabolic states as lipotypes.

Lipotypes mark specific cell transcriptional states

We performed single-cell RNA sequencing (scRNA-seq) on a total of 5652 dHFs. Uniform manifold approximation and projection (UMAP) embedding was computed on the gene expression profiles, and 17 cell clusters were identified by the Louvain algorithm (Fig. 4A). These 17 clusters were grouped into six categories related to different biological processes: proliferation, proinflammatory cytokine secretion (inflammatory), profibrotic secretion (fibrogenic), extracellular matrix remodeling (fibrolytic), and proangiogenic factor secretion (vascular) (Fig. 4B and fig. S5A). A further group represented bona fide basal-state fibroblasts (basal). We investigated the dynamic relationships among these categories using diffusion maps (*47*) and partition-based graph abstraction (PAGA), which estimate the trajectories and connectivity of the different components of a manifold (*48*). This analysis revealed that basal and proliferating categories were interconnected, whereas inflammatory, fibrogenic, and fibrolytic categories represented mutually alternative transcriptional cell configurations (Fig. 4, C and D).

Next, to link the expression-defined subtypes with those defined by sphingolipids, we isolated dHFs according to their lipotypes by FACS and performed bulk RNA sequencing on the different sorted samples. We isolated ChTxB$^+$, ShTxB2e$^+$, ShTxB1a$^+$/2e$^+$, and triple$^+$ cells (Fig. 4E and fig. S5B). Genes up-regulated in the different lipotypes were extracted and used to compute gene signature scores on the single-cell dataset. The four lipotype signatures mapped to distinct UMAP areas that corresponded to the major transcriptional categories (Fig. 4, F and G). Triple$^+$ cells corresponded to inflammatory, fibrolytic, and vascular fibroblasts; ShTxB1a$^+$/2e$^+$ and ShTxB2e$^+$ to proliferating cells and basal state fibroblasts; and ChTxB$^+$ to "fibrogenic" fibroblasts. This suggests that specific lipotypes are associated with prevalent cell states (Fig. 4H). To verify this finding, we costained dHFs with toxins and markers for the different clusters, which revealed a specific overlap between ChTxB and smooth muscle actin (encoded by the gene *ACTA2*, a fibrogenic marker)–positive cells and between ShTxB2e and laminin A (encoded by the gene *LMNA*, a basal and proliferative marker)–positive cells. Altogether, these observations indicate that lipotypes are markers of dHF cell states (Fig. 4, I and J).

Lipotypes mark specific dHF populations in vivo

Cell states of dHFs in vitro partly reflect populations of fibroblasts in the skin. Specifically, fibroblasts localized in the deeper dermal region (i.e., reticular fibroblasts) are endowed with fibrogenic activity (*49*), whereas those populating the more superficial region (i.e., papillary fibroblasts) have greater proliferative capability (*50*). Therefore, we derived transcriptional signatures for papillary and reticular dHFs from studies (*51*) and mapped them on our UMAP embedding (Fig. 5A). We found that the reticular signature largely overlaps with the fibrogenic UMAP region (also associated with ChTxB$^+$ signature). Conversely, the papillary signature overlaps with the basal and fibrolytic UMAP regions (also associated with ShTxB2e$^+$ and ShTxB1a$^+$/2e$^+$ signatures) (Fig. 5, A and B).

Accordingly, when we toxin-stained human skin biopsies, we observed that ChTxB$^+$ cells are preferentially found in the reticular dermal region, whereas ShTxB1a$^+$/2e$^+$ cells are prevalently found in the papillary dermal region (Fig. 5C). Counterstaining for the fibroblast marker vimentin and other dermal markers confirmed that dHFs are stained with different specificities by toxins (Fig. 5D and fig. S6). Keratinocytes, as recognized by the marker pankeratin, were primarily ShTxB1a$^+$/2e$^+$ and endothelial cells, as recognized by the marker CD31, were stained by all three toxins (Fig. 5E and fig. S6).

When skin is damaged, for example, from wounding or cancer lesions, dermal fibroblasts become activated and experience phenotypic interconversion (*52*). We stained three skin samples from individuals diagnosed with cutaneous squamous cell carcinoma (cSCC) with sphingolipid-binding toxins. In all three cases, recognizable cancer lesions were surrounded by cells prevalently stained by ChTxB (Fig. 5F and fig. S7, A and B). When counterstained with dermal markers, these ChTxB$^+$ cells were vimentin$^+$, suggesting that they are cancer-associated fibroblasts (CAFs) (fig. S7C).

CAFs can be effectively isolated from cancer tissues. We thus examined two pairs of CAFs and matched dHFs from cSCC and flanking unaffected areas from the same patients (fig. S7, D and E) for toxin analysis. In both cases, CAFs were predominantly ChTxB$^+$ and ShTxB1a$^-$/2e$^-$,

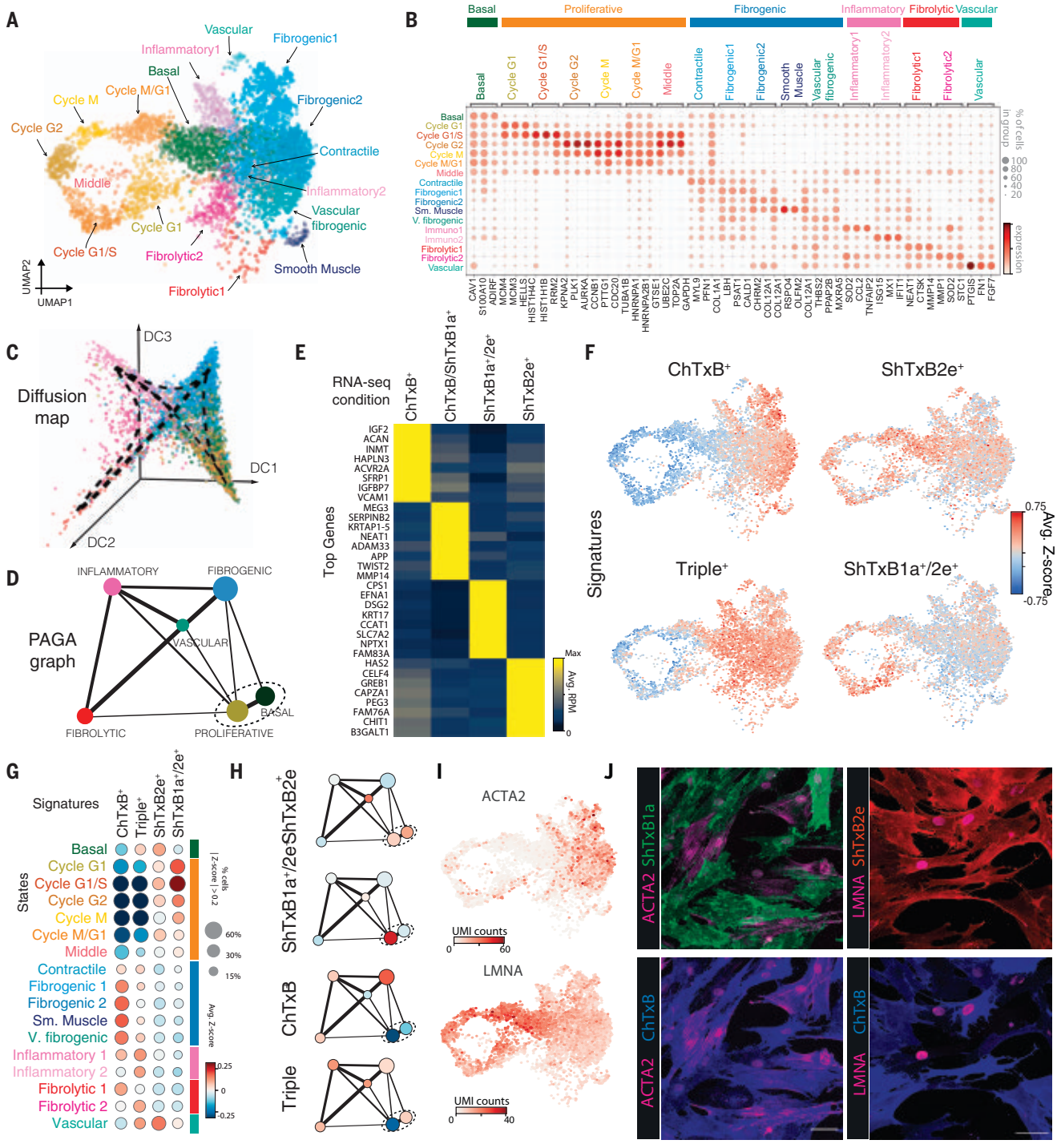


Fig. 4. Lipotype mapping to transcriptional cell states. (**A**) UMAP embedding analysis of scRNA-seq of 5652 individual dHFs colored by the assigned cluster. (**B**) Gene expression dot plot of cluster marker genes. Genes for each cluster were identified using the Wilcoxon rank-sum test. (**C**) Diffusion map visualization of single dHF cells from (A) highlighting the axes of transcriptional variation among the different cell states. (**D**) PAGA applied to scRNA-seq data of control dHFs. Nodes indicate cell type states, and the length of edges indicates the degree of similarity between states, with shorter edges corresponding to greater state similarity. (**E**) Heatmap reporting the average gene expression of enriched genes for each of the FACS-sorted lipotype populations (bulk RNA-seq data). For each lipotype, the top eight genes, ranked by fold change, are shown. (**F**) UMAP embedding colored by the different lipotype gene signature scores. The 250 top differentially expressed genes were used to calculate the signature score. (**G**) Dot plot colored by the average lipotype *z*-score of cells of the different clusters. Size of the dots represents the number of cells with magnitude of the score >0.35. (**H**) PAGA applied to scRNA-seq data of control dHFs and based on the ShTxB2e$^+$, ShTxB1a/2e$^+$, ChTxB$^+$, and triple$^+$ lipotype signatures. Nodes are positioned corresponding to cell states in (D). Color of nodes corresponds to *z*-score signatures, with a positive *z*-score (red) indicating a greater correspondence of the particular cell type to the particular lipotype state. Color bar is the same as in (G). (**I**) UMAP embedding of dHFs colored by the expression of the two canonical markers for fibrogenic (*ACTA2*) and basal (*LMNA*) cell states. (**J**) Confocal micrographs of cells stained with ShTxB1a (green), ShTxB2e (red), and ChTxB (blue) and counterstained by antibodies against ACTA2 and LMNA (magenta). Scale bar, 100 μm.

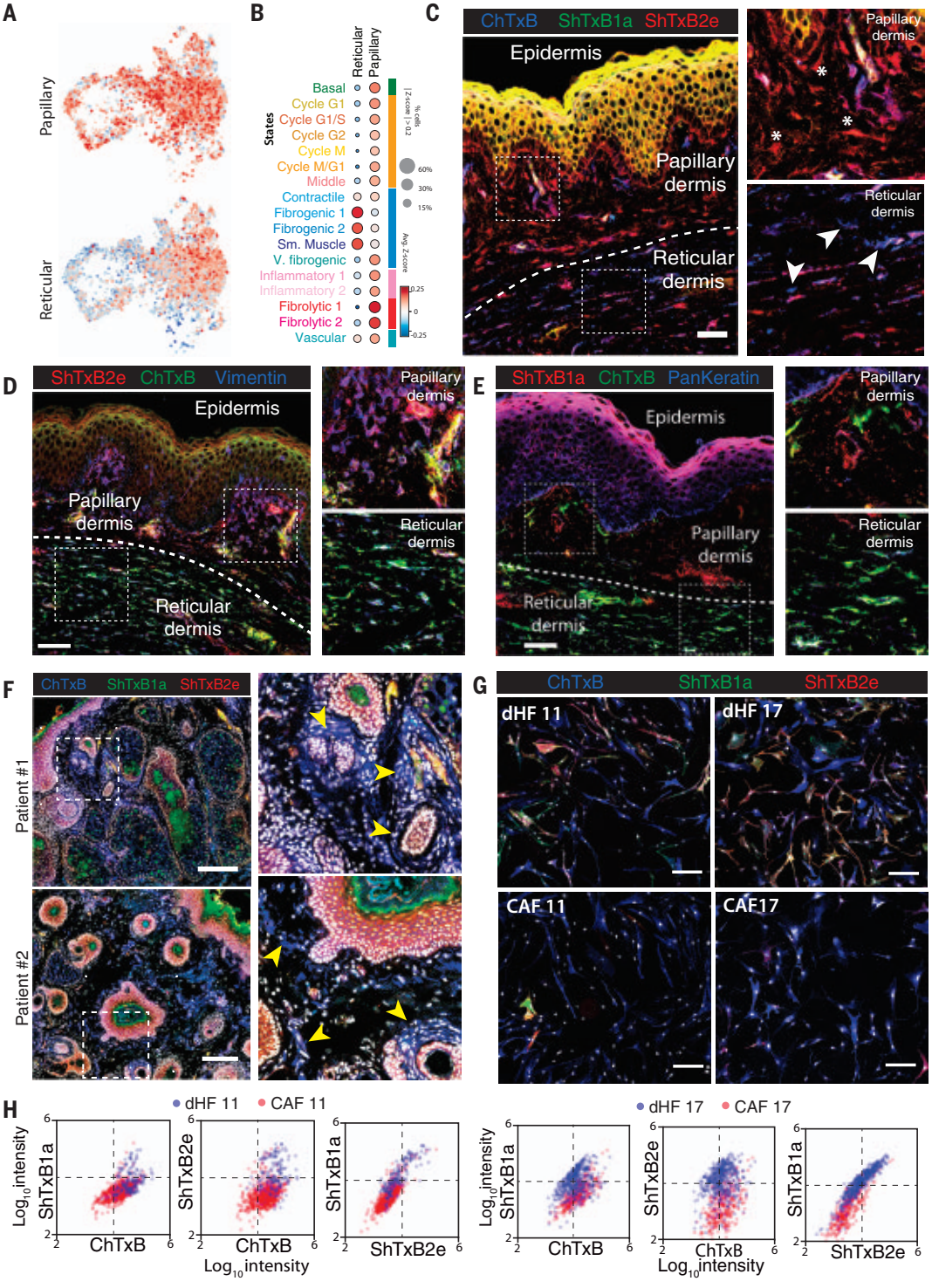


Fig. 5. Lipotypes define dHF population in the skin. (**A**) Papillary and reticular signatures overlaid onto the UMAP embedding. (**B**) Dot plot colored by the average reticular and papillary *z*-score of cells of the different clusters. Size of the dots represents the number of cells with magnitude of the score >0.35. (**C**) Confocal micrographs of human foreskin tissue section stained with bacterial toxins ShTxB1a (green), ShTxB2e (red), and ChTxB (blue). Insets show lipid staining in papillary (asterisks) and reticular (arrowheads) fibroblasts. Scale bar, 200 μm. The dotted line delineates the papillary-reticular dermal boundary. (**D**) Confocal micrographs of human foreskin tissue section stained with bacterial toxins ShTxB2e (red) and ChTxB (green) and vimentin (blue) as a fibroblast marker. Insets show staining in papillary and reticular layers. Scale bar, 200 μm. The dotted line delineates the papillary-reticular dermal boundary. (**E**) Confocal micrographs of human foreskin tissue section stained with bacterial toxins as in (D) and pankeratin (blue) as a keratinocyte marker. Insets show staining in papillary and reticular layers. Scale bar, 200 μm. (**F**) Confocal micrographs of human cSCC sections stained with bacterial toxins as in (C). Scale bar, 200 μm. Insets show tumor regions surrounded by ChTxB$^+$ fibroblasts (yellow arrowheads). (**G**) Confocal micrographs of CAFs and normal dHFs isolated from the same individuals, stained with bacterial toxins as in (C). Scale bar, 200 μm. (**H**) Scatter plots of fluorescence intensity values for each toxin comparing control dHFs (blue) and their corresponding CAFs (red).

compared with matched dHFs (Fig. 5, G and H). These data indicate that dHF lipotypes are reflected by fibroblast subtypes populating different dermal regions and are differently associated with skin cancers.

Sphingolipid composition influences cell states

We surveyed whether lipotypes are the result of cell state–specific transcriptional programs that involve lipid-metabolizing enzymes. Unexpectedly, when the expression of genes encoding sphingolipid enzymes and accessory factors was visualized on the UMAP embedding, none of them showed a cell state–specific localization (fig. S8A). To test this, we combined toxin staining and mRNA fluorescence in situ hybridization (FISH). We assayed the expression of *ST3GAL5* encoding GM3 synthase (GM3S) and *A4GALT* encoding Gb3 synthase (Gb3S) and their lipid products through toxins ChTxB and ShTxB1a within the same cells (Fig. 6A and fig. S8B). When toxin staining intensity was considered along with FISH counts, we observed either no or weak (R = 0.29) correlation of the two readouts, suggesting that single-cell sphingolipid composition is largely determined by posttranscriptional

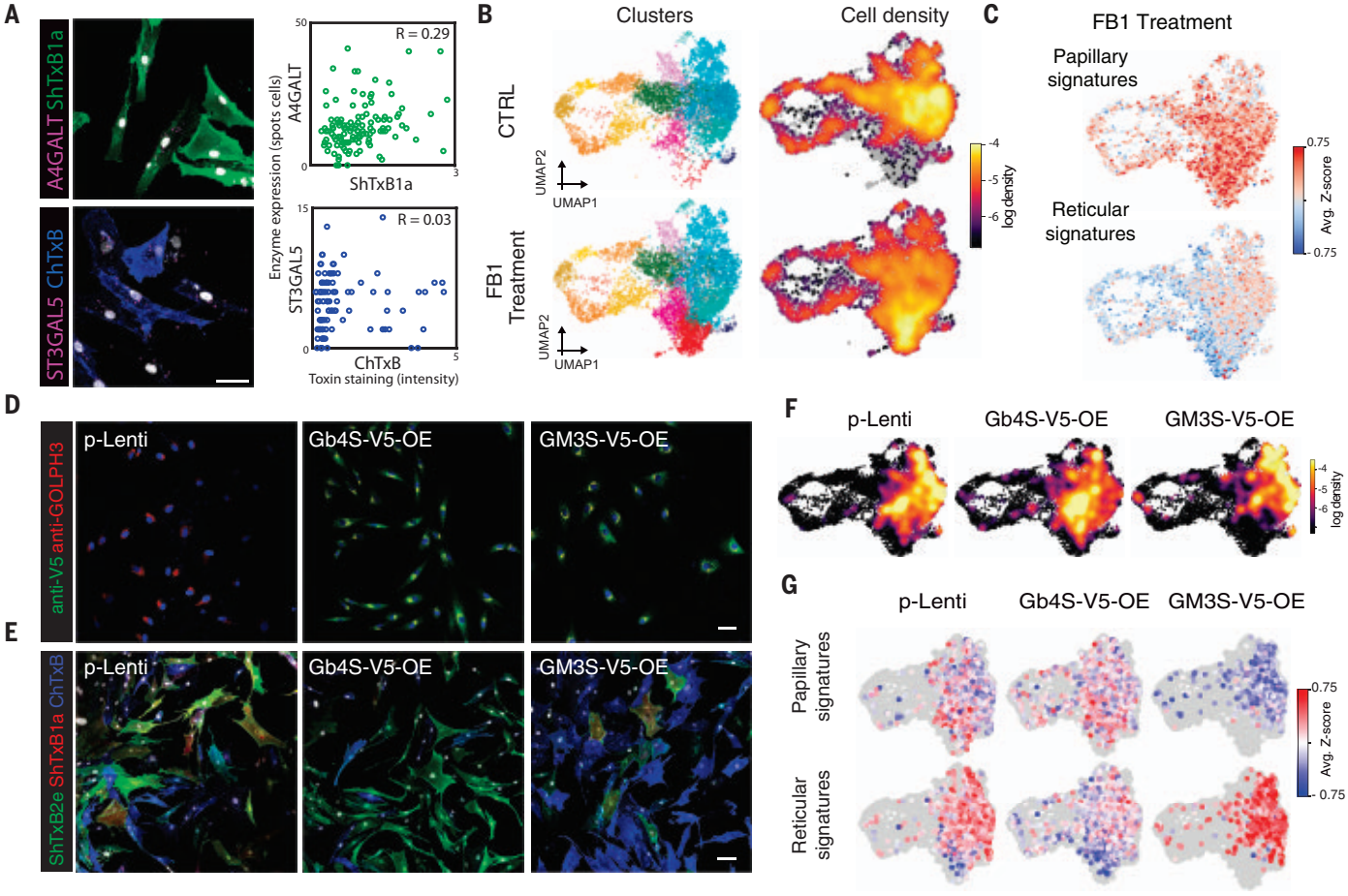


Fig. 6. Effect of sphingolipid perturbations on cell states. (**A**) Representative confocal micrographs of correlative mRNA-FISH/fluorescence toxin staining using *A4GALT* and *ST3GAL5* (magenta) probes and ShTxB1a (green) and ChTxB (blue). Nuclei were labeled with Hoechst (gray). Scale bar, 50 μm. Right, scatterplot showing the level of expression against toxin fluorescence. Pearson's correlation coefficient is indicated. Quantification on 120 and 96 individual cells for *A4GALT* and *ST3GAL5*, respectively. (**B**) Left, UMAP embedding of the scRNA-seq data for the control (5652 individual dHFs) and FB1-treated cells (6546 individual dHFs). Cells are colored by their assigned cluster. Right, density maps of control (CTRL) and FB1-treated cells mapped in the UMAP space. (**C**) Papillary and reticular gene signatures in the FB1-treated sample overlaid onto the UMAP embedding. (**D**) Confocal images of the overexpressing cell lines stained with antibodies against V5 protein tag (green) and GOLPH3 (red). Scale bar, 50 μm. (**E**) Confocal images of the overexpressing cell lines stained with the bacterial toxins ShTxB1a (red), ShTxB2e (green), and ChTxB (blue). Scale bar, 50 μm. (**F**) Cell density plot of single-cell expression profile of the OE dHF cells mapped by similarity onto the UMAP projection in (B). (**G**) Papillary and reticular signatures overlaid onto the UMAP embedding of the OE dHFs.

mechanisms. Therefore, lipotypes are associated with cell states, yet cell states are not endowed with transcriptional programs that would account for the lipotypes with which they are associated. This raises the question of whether lipotypes are causally upstream of cell states and if lipid composition influences cell-to-cell transcriptional heterogeneity.

To test this hypothesis, we treated dHFs with the Cer synthase inhibitor FB1, which blocks the production of sphingolipids (figs. S1F and S2B), and performed scRNA-seq. When FB1-treated dHFs were integrated in the same transcriptional embedding along with control cells, they displayed a different distribution across states (Fig. 6B and fig. S8C). FB1-treated cells were more frequently associated with fibrolytic (from 6% in control cells to 23% in FB1-treated cells) and vascular (from 0.6 to 1.3%) than fibrogenic (from 48 to 40%) and inflammatory (from 9 to 6%; Fig. 6B and fig. S8D) states. These changes correspond to an increased association of FB1-treated dHFs with "papillary" and decreased association with "reticular" fibroblast states (Fig. 6C).

FB1 treatment deprives cells of most sphingolipids (*43*), so this treatment does not inform on how the individual lipid species associated with the cell states influence signaling. We established dHF lines overexpressing either GM3S or Gb4S, two enzymes driving alternative sphingolipid-processing pathways (Fig. 6D). These overexpressing (OE) cells displayed the expected changes in sphingolipid composition, with GM3S-OE dHFs composed largely of ChTxB$^+$ cells and Gb4S-OE dHFs composed largely of ShTxB1a$^+$/2e$^+$ cells (Fig. 6E and fig. S8E).

GM3S-OE and Gb4S-OE lines were then analyzed by scRNA-seq to test the impact of lipotype change on cell state. GM3S-OE and Gb4S-OE dHFs populated two distinct transcriptional regions (Fig. 6F). Gb4S-OE dHFs were more associated with basal, inflammatory, and fibrolytic states, whereas GM3S-OE dHFs were for the major part in a fibrogenic state (88%) and almost never in inflammatory or fibrolytic states (fig. S8F). Gene expression analysis confirmed this transition: *COL12A1* and *VCAN* markers of fibrogenic state were significantly up-regulated in GM3S-OE cells and down-regulated in Gb4S-OE cells, whereas the fibrolytic and inflammatory markers *MMP-1*

and *CCL2* were up-regulated in Gb4S-OE cells and down-regulated in GM3S-OE cells (fig. S8G). Moreover, when reticular and papillary signatures were considered, we found that Gb4S-OE dHFs were clearly more associated with papillary states and GM3S-OE dHFs with reticular states (Fig. 6G).

Sphingolipids integrate into regulatory circuits involved in cell-state determination

The transcriptional changes triggered by FB1 involve targets of fibroblast growth factor 2 (FGF2) activation and of transforming growth factor–β (TGF-β) repression (*53*) (fig. S9A). FGF2 and TGF-β transcriptional programs are in fact largely antagonistic (*53*). Genes up-regulated by FGF2 and down-regulated by TGF-β (such as *MMP-1*) were preferentially expressed in the FB1-treated cells and in the fibrolytic-basal population, whereas genes up-regulated by TGF-β and down-regulated by FGF2 (such as *ACTA2*) were more expressed in control cells and in the fibrogenic population (fig. S9A). This was confirmed by quantitative polymerase chain reaction (qPCR) and immunofluorescence analysis on a panel of selected markers and extended to treatment with D-PDMP and LCS-KD (fig. S9, B to F), suggesting that global sphingolipid deprivation either promotes FGF2 or suppresses TGF-β signaling.

In dHFs challenged with increasing amounts of FGF2 or TGF-β, sphingolipid depletion did not inhibit fibroblast response to TGF-β, whereas it sensitized cells to FGF2 (fig. S9G). Moreover, genetic interruption of FGF signaling through the expression of a dominant-negative version of FGF receptor 1 (DNFGFR1) specifically blunted the transcriptional response to FB1 treatment (Fig. 7, A to C), indicating that transcriptional changes induced by sphingolipid deprivation require FGF signaling.

When FGF2 and TGF-β signatures were mapped onto Gb4S-OE and GM3S-OE dHF UMAPs, we observed that GM3S-OE fostered the TGF-β transcriptional program, whereas for Gb4S-OE, we revealed the opposite trend (fig. S9H). The effect on the FGF2 program was more difficult to observe because the expression signature dominates in actively proliferating cells (fig. S9H). Nonetheless, immunofluorescence experiments showed that although GM3S-OE dHFs were almost uniformly $ACTA2^{+}/MMP1^{-}$, Gb4S-OE dHFs displayed high MMP1 levels (Fig. 7, D and E). This effect was counteracted by the FGF signaling inhibitor infigratinib (fig. S9I), indicating again that transcriptional responses to changes in cellular sphingolipid composition require FGF signaling.

Moreover, stimulating Gb4S-OE and GM3S-OE dHFs with FGF2 resulted in increased and decreased responses, respectively (fig. S9, J and K), and exogenous administration of GM1 to FB1-treated cells (fig. S9L) specifically counteracted MMP1 induction (fig. S9M). This suggests that GM1 and Gb3/Gb4 have opposite modulatory effects on FGF2 signaling. We thus challenged dHFs with FGF2 and monitored the immediate single-cell signaling response by following ERK phosphorylation (*54*) as a function of the cell lipotype. In our conditions, FGF2-induced ERK phosphorylation was maximal after 5 min of stimulation (Fig. 7F and fig. S10A). At this time point, $ShTxB1a^{+}/2e^{+}$ cells displayed a consistently stronger response to FGF2 than $ChTxB^{+}$ cells from the same cell culture dish (Fig. 7G and fig S10, B and C), indicating that dHFs exposing Gb3 and Gb4 at their cell surfaces are more susceptible to FGF pathway activation than those exposing GM1.

Unexpectedly, toxin staining analysis of dHFs in which the FGF2 pathway was blocked either genetically or pharmacologically showed a transition of the dHFs to a $ChTxB^{+}$ state (Fig. 7H and fig. S10, D and E). Along similar lines, FGF2 stimulation induced an increase in the $ShtxB1a^{+}/2e^{+}$ cell population with a concurrent decrease of $ChTxB^{+}$ cells (fig. S10, D and E). This effect was largely nontranscriptional, because in DNFGFR1 dHFs, the production of the Gb3 was reduced as assessed by metabolic labeling (fig. S10F), yet the expression of the genes that encode sphingolipid synthetic enzymes was not modulated (fig. S10G).

Thus, sphingolipids modulate FGF2 signaling, with Gb3/Gb4 acting as positive regulators and GM1 as a negative regulator. In turn, FGF2 signaling counteracts GM1 production by sustaining the alternative metabolic pathway leading to the production of Gb3 and Gb4 (Fig. 7I).

Discussion

Here, we investigated whether and how lipid metabolism affects cell identity by exploring the dHF heterogeneity (*7*, *14*) that results from their plastic interconversion across cell states (*6*, *55*–*57*).

Our observations constitute an example of how cell-to-cell lipid heterogeneity can diversify the processing of extracellular signals and promote cellular responses (*22*). The phenomenon that we describe can be considered an instance of cellular contextual decision-making whereby individual cells route to alternative fates by processing external inputs in the context of their internal states (*58*).

Furthermore, considering both the ubiquity of lipids and their structural diversity, we expect to find other cell types exploiting regulatory strategies analogous to the one that we discovered. By extension, one can hypothesize that lipidome remodeling participates in tissue patterning and organogenesis. If this is correct, then lipid-defined cell states analogous to the lipotypes described here could be involved in developmental symmetry-breaking events and organogenesis (*4*). Indeed, our finding that lipotypes are spatially segregated to different dermal layers in human skin architecture supports this hypothesis.

A limitation of our study is the inability to address lipid and transcriptional trajectories live and in single cells. Although challenging to obtain, time-resolved data have the potential to clarify how lipid metabolic fluxes evolve during cell-state transitions (*59*). We envision that emerging tools such as chemically synthesized lipid probes and live-omics profiling will enable such experiments (*60*, *61*).

In conclusion, by exploiting the potential of space-resolved nontranscriptional single-cell omics, we provide evidence for cell-to-cell heterogeneous lipid metabolism playing an instructive role in the self-organization of multicellular systems.

Methods summary

Human fibroblasts obtained from the dermis of discarded skin samples of circumcised, 1- to 5-year-old healthy males were used for MALDI-MSI analyses. Specifically, samples were analyzed using AP-SMALDI10 or AP-SMALDI5 AF systems using 5- or 7-μm spatial resolution in positive-ion mode in the mass range m/z 400 to 1600. Mass images (n = 296) were then generated and lipids annotated by using a combination of databases, ESI-LC/MS (*62*), and MRM confirmation. To assess lipid variability, single-pixel analysis was performed on the 296 mass images. PCA analysis was performed and the absolute values of the PCA loadings were then used to identify the lipids with the most variance of each single component. Single dHFs were further manually segmented, and raw abundance data for each scan and each pixel in a cell were exported. Normalized lipid count values were used to determine the CV. Pearson's R was used to evaluate lipid and cell covariation.

For lipotype determination and feature extraction, including fluorescence intensities, area, eccentricity, shape complexity, and local cell density, cells were stained with fluorescently labeled B-subunit toxins or primary and secondary antibodies. Cells were then analyzed by confocal microscopy and segmented using Cellpose. Time-lapse imaging coupled with toxin end-point staining was performed to assess the dynamics of lipotype configuration (*63*, *64*). The lineage information extracted from the time-lapse imaging and the cell state from the end-point staining were used to perform a sister-state frequency analysis and to fit the estimation framework CELLMA (*37*, *65*).

To evaluate the association between transcriptional states and lipotypes, scRNA-seq on

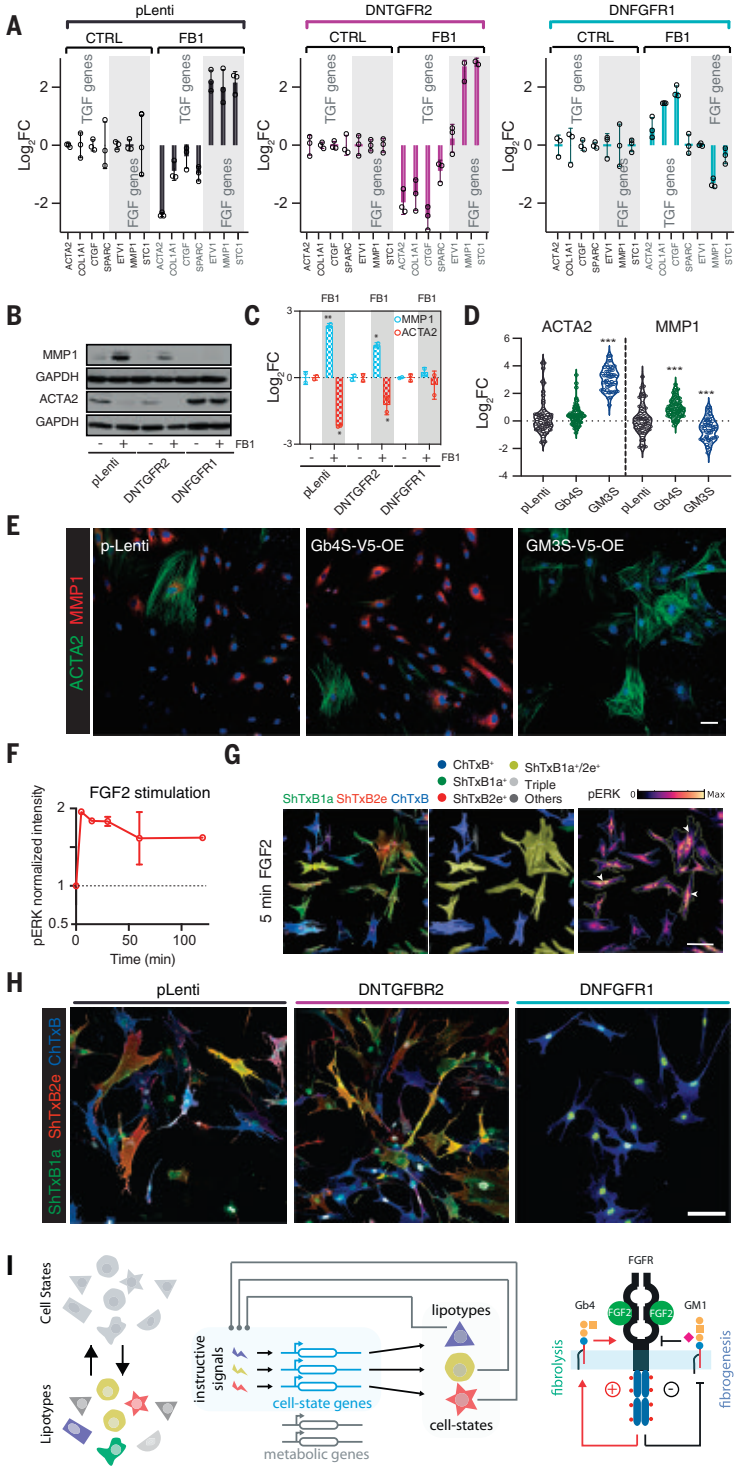


Fig. 7. Effect of sphingolipid perturbation on FGF signaling. (**A**) Barplots of qPCR quantifying the mRNA levels of TGF-β and FGF target genes in CTRL, DNTGFR2, and DNFGFR1 cells treated with 25 μM FB1 for 6 days. Data are shown as log_2-fold change over untreated cells (n = 3; data are shown as means ± SD). (**B** and **C**) Western blot and quantification of cells treated as in (A). Data were normalized against GAPDH (n = 2; data are means ± SD; *P < 0.05, **P < 0.01, Student's t test). (**D**) Cells were stained with ACTA2 and MMP1, and normalized intensity values were extracted for quantification. Data were scaled to the median and are shown as the log_2-fold change over control in an individual cell (CTRL, n = 58; GM3S-OE, n = 49; Gb4S-OE, n = 81; ***P < 0.001, ordinary one-way ANOVA). (**E**) Representative confocal images of OE dHFs stained with antibodies against ACTA2 (green) and MMP1 (red). Scale bar, 100 μm. (**F**) Plots indicating normalized intensity values of pERK protein in cells serum starved and then treated with 5 ng/ml of FGF2 for different times as determined by densitometry. (**G**) Cells treated for 5 min with 5 ng/ml of FGF2 as in (F) were stained with the bacterial toxins ShTxB1a (green), ShTxB2e (red), and ChTxB (blue) (left panel) and pERK (gradient) (right panel). Representative confocal micrographs and cell segmentations according to lipotype are shown. Scale bar, 100 μm. (**H**) Confocal images of pLenti, DNTGFBR2, and DNFGFR1 cells stained with the bacterial toxins ShTxB1a (green), ShTxB2e (red), ChTxB (blue), and Hoechst for nuclei. Scale bar, 100 μm. (**I**) Schematic representation of the model for the role of lipotypes in cell-state determination. Left panel, lipotypes corresponding to dHF cell states. Middle panel, cell states and lipotypes determined by signaling pathways that are in turn influenced by the lipid composition of individual cells. Right panel, FGF2 binds to FGFR, leading to the prevalent production of Gb3/Gb4 over GM1. To close the circuit, GM1 negatively regulates FGFR, whereas Gb3 and Gb4 activate FGFR in a positive feedback loop. This is a bistable system in which cells can be $Gb3^+$ or $Gb4^+$, leading to a more fibrolytic state, or $GM1^+$, leading to a more fibrogenic state.

dHFs and bulk RNA-seq on five cell populations isolated by FACS according to their lipid composition were performed. Skin tissue sections isolated from healthy individuals or from patients diagnosed with cSCC were used to evaluate the sphingolipid composition in vivo after toxin staining and confocal microscopy. To assess the influence of sphingolipid composition on cell state, dHFs treated with the Cer synthase inhibitor FB1 or lentiviral stable cell lines overexpressing GSL-synthesizing enzymes were analyzed by scRNA-seq.

Full materials and methods are available as supplementary materials (*37*).

REFERENCES AND NOTES

1. D. Wang, S. Bodovitz, Single cell analysis: The new frontier in 'omics'. *Trends Biotechnol.* **28**, 281–290 (2010). doi: 10.1016/j.tibtech.2010.03.002; pmid: 20434785
2. O. Symmons, A. Raj, What's luck got to do with it: Single cells, multiple fates, and biological nondeterminism. *Mol. Cell* **62**, 788–802 (2016). doi: 10.1016/j.molcel.2016.05.023; pmid: 27259209
3. J. P. Junker, A. van Oudenaarden, Every cell is special: Genome-wide studies add a new dimension to single-cell biology. *Cell* **157**, 8–11 (2014). doi: 10.1016/j.cell.2014.02.010; pmid: 24679522
4. Q. Chen, J. Shi, Y. Tao, M. Zernicka-Goetz, Tracing the origin of heterogeneity and symmetry breaking in the early mammalian embryo. *Nat. Commun.* **9**, 1819 (2018). doi: 10.1038/s41467-018-04155-2; pmid: 29739935
5. M. Zinner, I. Lukonin, P. Liberali, Design principles of tissue organisation: How single cells coordinate across scales. *Curr. Opin. Cell Biol.* **67**, 37–45 (2020). doi: 10.1016/j.ceb.2020.07.004; pmid: 32889170

6. R. R. Driskell, F. M. Watt, Understanding fibroblast heterogeneity in the skin. *Trends Cell Biol.* **25**, 92–99 (2015). doi: 10.1016/j.tcb.2014.10.001; pmid: 25455110
7. J. M. Sorrell, A. I. Caplan, Fibroblast heterogeneity: More than skin deep. *J. Cell Sci.* **117**, 667–675 (2004). doi: 10.1242/jcs.01005; pmid: 14754903
8. F. M. Watt, H. Fujiwara, Cell-extracellular matrix interactions in normal and diseased skin. *Cold Spring Harb. Perspect. Biol.* **3**, a005124 (2011). doi: 10.1101/cshperspect.a005124; pmid: 21441589
9. C. Philippeos *et al*., Spatial and single-cell transcriptional profiling identifies functionally distinct human dermal fibroblast subpopulations. *J. Invest. Dermatol.* **138**, 811–825 (2018). doi: 10.1016/j.jid.2018.01.016; pmid: 29391249
10. E. Rognoni *et al*., Fibroblast state switching orchestrates dermal maturation and wound healing. *Mol. Syst. Biol.* **14**, e8174 (2018). doi: 10.15252/msb.20178174; pmid: 30158243
11. A. M. Ascensión, S. Fuertes-Álvarez, O. Ibañez-Solé, A. Izeta, M. J. Araúzo-Bravo, Human dermal fibroblast subpopulations are conserved across single-cell RNA sequencing studies. *J. Invest. Dermatol.* **141**, 1735–1744.e35 (2021). doi: 10.1016/j.jid.2020.11.028; pmid: 33385399
12. M. Adler *et al*., Principles of cell circuits for tissue repair and fibrosis. *iScience* **23**, 100841 (2020). doi: 10.1016/j.isci.2020.100841; pmid: 32058955
13. E. Rognoni, F. M. Watt, Skin cell heterogeneity in development, wound healing, and cancer. *Trends Cell Biol.* **28**, 709–722 (2018). doi: 10.1016/j.tcb.2018.05.002; pmid: 29807713
14. M. F. Griffin, H. E. desJardins-Park, S. Mascharak, M. R. Borrelli, M. T. Longaker, Understanding the impact of fibroblast heterogeneity on skin fibrosis. *Dis. Model. Mech.* **13**, dmm044164 (2020). doi: 10.1242/dmm.044164; pmid: 32541065
15. G. Biffi, D. A. Tuveson, Diversity and biology of cancer-associated fibroblasts. *Physiol. Rev.* **101**, 147–176 (2021). doi: 10.1152/physrev.00048.2019; pmid: 32466724
16. C. F. Guerrero-Juarez *et al*., Single-cell analysis reveals fibroblast heterogeneity and myeloid-derived adipocyte progenitors in murine skin wounds. *Nat. Commun.* **10**, 650 (2019). doi: 10.1038/s41467-018-08247-x; pmid: 30737373
17. Y. Rinkevich *et al*., Skin fibrosis. Identification and isolation of a dermal lineage with intrinsic fibrogenic potential. *Science* **348**, aaa2151 (2015). doi: 10.1126/science.aaa2151; pmid: 25883361
18. E. Rognoni *et al*., Inhibition of β-catenin signalling in dermal fibroblasts enhances hair follicle regeneration during wound healing. *Development* **143**, 2522–2535 (2016). doi: 10.1242/dev.131797; pmid: 27287810
19. T. S. Cliff, S. Dalton, Metabolic switching and cell fate decisions: Implications for pluripotency, reprogramming and development. *Curr. Opin. Genet. Dev.* **46**, 44–49 (2017). doi: 10.1016/j.gde.2017.06.008; pmid: 28662447
20. D. Russo, L. Capolupo, J. S. Loomba, L. Sticco, G. D'Angelo, Glycosphingolipid metabolism in cell fate specification. *J. Cell Sci.* **131**, jcs219204 (2018). doi: 10.1242/jcs.219204; pmid: 30559216
21. B. Snijder *et al*., Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature* **461**, 520–523 (2009). doi: 10.1038/nature08282; pmid: 19710653
22. M. Frechin *et al*., Cell-intrinsic adaptation of lipid composition to local crowding drives social behaviour. *Nature* **523**, 88–91 (2015). doi: 10.1038/nature14429; pmid: 26009010
23. M. Denz *et al*., Cell cycle dependent changes in the plasma membrane organization of mammalian cells. *Biochim. Biophys. Acta Biomembr.* **1859**, 350–359 (2017). doi: 10.1016/j.bbamem.2016.12.004; pmid: 27993567
24. D. Russo *et al*., Glycosphingolipid metabolic reprogramming drives neural differentiation. *EMBO J.* **37**, e97674 (2018). doi: 10.15252/embj.201797674; pmid: 29282205
25. M. Vietri Rudan, A. Mishra, C. Klose, U. S. Eggert, F. M. Watt, Human epidermal stem cell differentiation is modulated by specific lipid subspecies. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 22173–22182 (2020). doi: 10.1073/pnas.2011310117; pmid: 32843345
26. R. Zenobi, Single-cell metabolomics: Analytical and biological perspectives. *Science* **342**, 1243259 (2013). doi: 10.1126/science.1243259; pmid: 24311695
27. C. Thiele, K. Wunderling, P. Leyendecker, Multiplexed and single cell tracing of lipid metabolism. *Nat. Methods* **16**, 1123–1130 (2019). doi: 10.1038/s41592-019-0593-6; pmid: 31611692
28. L. Rappez *et al*., SpaceM reveals metabolic states of single cells. *Nat. Methods* **18**, 799–805 (2021). doi: 10.1038/s41592-021-01198-0; pmid: 34226721
29. J. L. Norris, R. M. Caprioli, Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. *Chem. Rev.* **113**, 2309–2342 (2013). doi: 10.1021/cr3004295; pmid: 23394164
30. Y. Schober, S. Guenther, B. Spengler, A. Römpp, Single cell matrix-assisted laser desorption/ionization mass spectrometry imaging. *Anal. Chem.* **84**, 6293–6297 (2012). doi: 10.1021/ac301337h; pmid: 22816738
31. M. E. Dueñas, J. J. Essner, Y. J. Lee, 3D MALDI mass spectrometry imaging of a single cell: Spatial mapping of lipids in the embryonic development of zebrafish. *Sci. Rep.* **7**, 14946 (2017). doi: 10.1038/s41598-017-14949-x; pmid: 29097697
32. A. Zavalin *et al*., Direct imaging of single cells and tissue at sub-cellular spatial resolution using transmission geometry MALDI MS. *J. Mass Spectrom.* **47**, 1473–1481 (2012). doi: 10.1002/jms.3108; pmid: 23147824
33. M. Kompauer, S. Heiles, B. Spengler, Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4-μm lateral resolution. *Nat. Methods* **14**, 90–96 (2017). doi: 10.1038/nmeth.4071; pmid: 27842060
34. M. Niehaus, J. Soltwisch, M. E. Belov, K. Dreisewerd, Transmission-mode MALDI-2 mass spectrometry imaging of cells and tissues at subcellular resolution. *Nat. Methods* **16**, 925–931 (2019). doi: 10.1038/s41592-019-0536-2; pmid: 31451764
35. E. Sugiyama, I. Yao, M. Setou, Visualization of local phosphatidylcholine synthesis within hippocampal neurons using a compartmentalized culture system and imaging mass spectrometry. *Biochem. Biophys. Res. Commun.* **495**, 1048–1054 (2018). doi: 10.1016/j.bbrc.2017.11.108; pmid: 29162450
36. D. R. Bhandari, G. Coliva, M. Fedorova, B. Spengler, "Single cell analysis by high-resolution atmospheric-pressure MALDI MS imaging," in: *Single Cell Metabolism: Methods and Protocols*, B. Shrestha, Ed. (Springer, 2020); pp. 103–111.
37. Materials and methods are available as supplementary materials.
38. J. C. M. Holthuis, A. K. Menon, Lipid landscapes and pipelines in membrane homeostasis. *Nature* **510**, 48–57 (2014). doi: 10.1038/nature13474; pmid: 24899304
39. L. Pelkmans, Cell Biology. Using cell-to-cell variability—A new era in molecular biology. *Science* **336**, 425–426 (2012). doi: 10.1126/science.1222161; pmid: 22539709
40. M. Jacewicz, H. Clausen, E. Nudelman, A. Donohue-Rolfe, G. T. Keusch, Pathogenesis of shigella diarrhea. XI. Isolation of a shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *J. Exp. Med.* **163**, 1391–1404 (1986). doi: 10.1084/jem.163.6.1391; pmid: 3519828
41. J. Müthing, C. H. Schweppe, H. Karch, A. W. Friedrich, Shiga toxins, glycosphingolipid diversity, and endothelial cell injury. *Thromb. Haemost.* **101**, 252–264 (2009). doi: 10.1160/TH08-05-0317; pmid: 19190807
42. S. van Heyningen; Heyningen S Van, Cholera toxin: Interaction of subunits with ganglioside GM1. *Science* **183**, 656–657 (1974). doi: 10.1126/science.183.4125.656; pmid: 4810267
43. E. Wang, W. P. Norred, C. W. Bacon, R. T. Riley, A. H. Merrill Jr., Inhibition of sphingolipid biosynthesis by fumonisins. Implications for diseases associated with Fusarium moniliforme. *J. Biol. Chem.* **266**, 14486–14490 (1991). doi: 10.1016/S0021-9258(18)98712-0; pmid: 1860857
44. J. Inokuchi, K. Momosaki, H. Shimeno, A. Nagamatsu, N. S. Radin, Effects of D-threo-PDMP, an inhibitor of glucosylceramide synthetase, on expression of cell surface glycolipid antigen and binding to adhesive proteins by B16 melanoma cells. *J. Cell. Physiol.* **141**, 573–583 (1989). doi: 10.1002/jcp.1041410316; pmid: 2531751
45. P. Liberali, B. Snijder, L. Pelkmans, A hierarchical map of regulatory genetic interactions in membrane trafficking. *Cell* **157**, 1473–1487 (2014). doi: 10.1016/j.cell.2014.04.029; pmid: 24906158
46. Y. A. Hannun, L. M. Obeid, Sphingolipids and their metabolism in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 175–191 (2018). doi: 10.1038/nrm.2017.107; pmid: 29165427
47. L. Haghverdi, F. Buettner, F. J. Theis, Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* **31**, 2989–2998 (2015). doi: 10.1093/bioinformatics/btv325; pmid: 26002886
48. F. A. Wolf *et al*., PAGA: Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019). doi: 10.1186/s13059-019-1663-x; pmid: 30890159
49. R. R. Driskell *et al*., Distinct fibroblast lineages determine dermal architecture in skin development and repair. *Nature* **504**, 277–281 (2013). doi: 10.1038/nature12783; pmid: 24336287
50. R. A. Harper, G. Grove, Human skin fibroblasts derived from papillary and reticular dermis: Differences in growth potential in vitro. *Science* **204**, 526–527 (1979). doi: 10.1126/science.432659; pmid: 432659
51. D. G. Janson, G. Saintigny, A. van Adrichem, C. Mahé, A. El Ghalbzouri, Different gene expression patterns in human papillary and reticular fibroblasts. *J. Invest. Dermatol.* **132**, 2565–2572 (2012). doi: 10.1038/jid.2012.192; pmid: 22696053
52. H. Y. Chang *et al*., Gene expression signature of fibroblast serum response predicts human cancer progression: Similarities between tumors and wounds. *PLOS Biol.* **2**, E7 (2004). doi: 10.1371/journal.pbio.0020007; pmid: 14737219
53. P. Bordignon *et al*., Dualism of FGF and TGF-β signaling in heterogeneous cancer-associated fibroblast activation with ETV1 as a critical determinant. *Cell Rep.* **28**, 2358–2372.e6 (2019). doi: 10.1016/j.celrep.2019.07.092; pmid: 31461652
54. L. Gerosa *et al*., Receptor-driven ERK pulses reconfigure MAPK signaling and enable persistence of drug-adapted BRAF-mutant melanoma cells. *Cell Syst.* **11**, 478–494.e9 (2020). doi: 10.1016/j.cels.2020.10.002; pmid: 33113355
55. S. Abbasi *et al*., Distinct regulatory programs control the latent regenerative potential of dermal fibroblasts during wound healing. *Cell Stem Cell* **27**, 396–412.e6 (2020). doi: 10.1016/j.stem.2020.07.008; pmid: 32755548
56. M. V. Plikus *et al*., Regeneration of fat cells from myofibroblasts during wound healing. *Science* **355**, 748–752 (2017). doi: 10.1126/science.aai8792; pmid: 28059714
57. D. Jiang, Y. Rinkevich, Scars or regeneration? Dermal fibroblasts as drivers of diverse skin wound responses. *Int. J. Mol. Sci.* **21**, 617 (2020). doi: 10.3390/ijms21020617; pmid: 31963533
58. B. A. Kramer, L. Pelkmans, Cellular state determines the multimodal signaling response of single cells. *bioRxiv*. 2019. p. 2019.12.18.880930. doi: 10.1101/2019.12.18.880930
59. A. R. Lederer, G. La Manno, The emergence and promise of single-cell temporal-omics approaches. *Curr. Opin. Biotechnol.* **63**, 70–78 (2020). doi: 10.1016/j.copbio.2019.12.005; pmid: 31918114
60. W. Chen *et al*., Genome-wide molecular recording using Live-seq. *bioRxiv*. 2021. p. 2021.03.24.436752. doi: 10.1101/2021.03.24.436752
61. M. Schuhmacher *et al*., Live-cell lipid biochemistry reveals a role of diacylglycerol side-chain composition for cellular lipid dynamics and protein affinities. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 7729–7738 (2020). doi: 10.1073/pnas.1912684117; pmid: 32213584
62. L. Capolupo, "Untargeted lipidomics of primary human skin fibroblasts" (NIH Metabolomics Workbench, 2020); https://www.metabolomicsworkbench.org/data/DRCCMetadata.php?Mode=Project&ProjectID=PR001087.
63. L. Capolupo, O. Burri, R. Guiet, Cellpose model for digital phase contrast images for: Sphingolipids control dermal fibroblast heterogeneity, Zenodo (2022); https://zenodo.org/record/6023317#.YjTffurMKUk.
64. L. Capolupo, Digital phase contrast on primary dermal human fibroblasts cells for: Sphingolipids control dermal fibroblast heterogeneity, Zenodo (2022); https://zenodo.org/record/5996883#.YjTfourMKUk.
65. L. Capolupo, I. Khven, A. R. Lederer, G. La Manno, G. D'Angelo, Notebooks for analysis of single-cell RNA sequencing datasets and CELLMA lipotype state transition inference modeling for: Sphingolipids control dermal fibroblast heterogeneity, Zenodo (2022); https://zenodo.org/record/6245944#.YjTf7erMKUk.

ACKNOWLEDGMENTS

We thank B. Deplancke, P. Seetharaman, D. Russo, C. Dibner, Y. Hannun, C. Luberto, and S. Linnarsson for critically reading the manuscript; H. H. Schede for the help with MALDI-MSI image analysis; and L. Talamanca for the comments on mathematical notation. **Funding:** G.L.M. and I.K. were supported by the Swiss National Science Foundation (grants CRSK-3_190495 and PZ00P3_193445). G.L.M. was supported by the School of Life Sciences, EPFL. G.D. acknowledges financial support from the Swiss Cancer League (grant KFS-4999-02-2020), the EPFL institutional fund, the Kristian Gerhard Jebsen Foundation, and the Swiss National Science Foundation (SNSF grant 310030_184926). L.M. and G.P.D. were supported by the Swiss National Science Foundation (grant 310030B_176404), the National Institutes of Health (NIH grant R01AR039190; the content does not necessarily represent the official views of the NIH), and the European Union's Horizon 2020 research and innovation program (Marie Skłodowska-Curie grant 859860). **Author contributions:**

L.C. developed the idea, conducted the experiments, and wrote the manuscript. I.K. participated in the development of the idea, conducted image and scRNA-seq analyses, and performed the scRNA-seq experiments. A.R.L. conceptualized and contributed to implementing the CELLMA model. L.M. supported L.C. in experiments on DN fibroblast lines. G.G. generated the GM3S and Gb4S OE fibroblast lines. S.H. provided technical assistance. F.R. performed initial lipid data analysis. J.P.M. performed targeted and untargeted lipidomics. A.P.B. and S.R.E. assisted with MALDI-MSI. R.G. performed FISH analysis. O.B. helped with live imaging experiments. J.M. and J.D. provided ShTxB2e toxin. K.H. contributed to discussions and provided reagents. F.K. and M.G. provided skin tissue sections. D.R.B. and B.S. performed AP-SMALDI10 and AP-SMALDI5 AF experiments. R.M.A.H. provided assistance with MALDI-MSI data treatment. G.P.D. contributed to discussions and provided reagents for dHF heterogeneity. G.L.M. and G.D.A. developed the idea, designed and supervised the entire project, analyzed the data, and wrote the manuscript. **Competing interests:** B.S is a consultant and D.R.B. is a part-time employee of TransMIT GmbH, Giessen, Germany. The authors declare no competing financial interests. **Data availability:** All data are available in the main text and supplementary materials and at (*62*–*65*). Further information and requests for resources and reagents can be directed to the corresponding authors. Any requests for toxins and antibodies can be directed to the corresponding authors or to the contact listed in table S4. Sequencing data are available on the Gene Expression Omnibus database under accession no. GSE167209. Lipid images are available at https://metaspace2020.eu/project/capolupo-2022.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abh1623
Materials and Methods
Supplementary Note 1
Figs. S1 to S10
Tables S1 to S6
References (*66*–*78*)
Movie S1
MDAR Reproducibility Checklist

20 February 2021; resubmitted 11 November 2021
Accepted 7 March 2022
10.1126/science.abh1623

RESEARCH ARTICLE SUMMARY

NEUROIMMUNOLOGY

Bacterial sensing via neuronal Nod2 regulates appetite and body temperature

Ilana Gabanyi*, Gabriel Lepousez, Richard Wheeler, Alba Vieites-Prado, Antoine Nissant, Sébastien Wagner, Carine Moigneu, Sophie Dulauroy, Samia Hicham, Bernadette Polomack, Florine Verny, Philip Rosenstiel, Nicolas Renier, Ivo Gomperts Boneca, Gérard Eberl*†, Pierre-Marie Lledo*†

INTRODUCTION: Compounds released by the microbiota are found in the bloodstream and can modulate physiological processes in the host, such as immunity, metabolism, and brain functions. Microbial metabolites, including short-chain fatty acids and tryptophan derivatives, regulate many processes through receptors that are widely expressed. However, the structural components of microbes are detected by pattern recognition receptors (PRRs) that signal the presence of viruses, bacteria, or fungi on mucosal surfaces, in tissues, and in cells. Bacterial components have been found to modulate brain activity, and PRRs are associated with brain disorders. Whether brain neurons can directly sense bacterial components and whether bacteria can regulate physiological processes through regulation of brain neurons remains to be demonstrated.

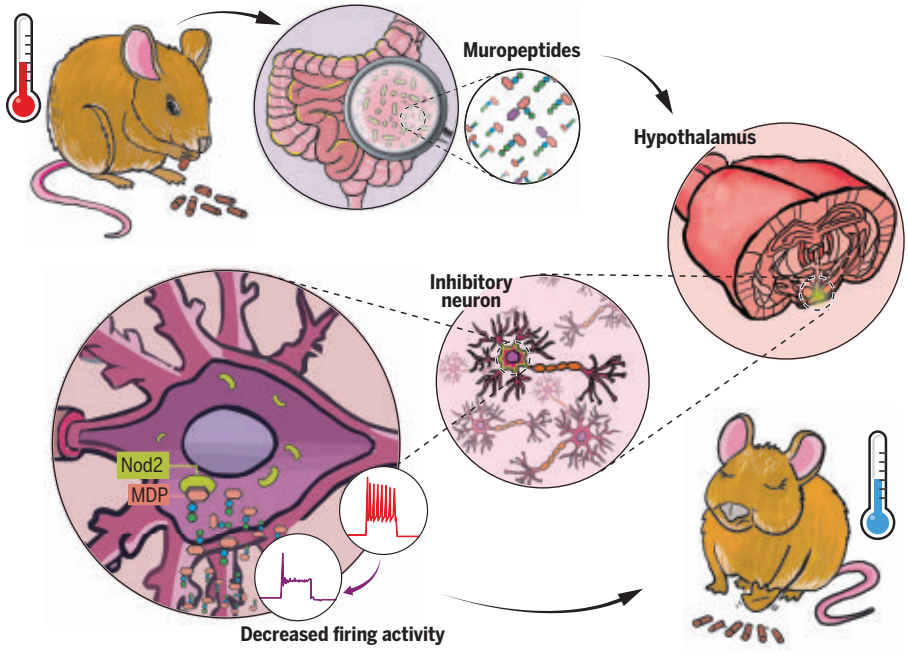
RATIONALE: Mice that lack neuronal expression of Nod2—a PRR recognizing fragments of the bacterial cell wall termed muropeptides—develop alterations in food intake, nesting behavior, and body temperature control. We used brain imaging to identify regions affected by the oral administration of muropeptides and measured the modulation of neuronal activity by muropeptides. We also developed mice that lack Nod2 expression in subsets of neurons and in regions of the hypothalamus that regulate feeding behavior and body temperature, so as to assess the impact of the gut-brain axis on the regulation of host metabolism. Finally, we used patch-clamp recordings to assess whether neurons directly respond to muropeptides.

RESULTS: Using reporter mice and in situ hybridization techniques, we found that Nod2 was expressed in neurons of several brain regions including the hypothalamus. Older female mice lacking expression of Nod2 in inhibitory γ-aminobutyric acid transporter–positive (GABAergic) neurons ate more and consequently gained more weight than normal mice. Oral administration of muramyl dipeptide (MDP), a muropeptide ligand of Nod2, reduced feeding but only when activating Nod2 in GABAergic neurons. These mice also showed a reduced propensity to build nests, a behavioral trait related to heat conservation, as well as reduced temperature regulation in response to the circadian rhythm, fasting, and adrenergic stimulation.

MDP, administered orally or as muropeptides produced by intestinal bacteria, reached the brain and regulated neurons in diverse brain areas of older female mice, including the arcuate nucleus of the hypothalamus, which is involved in the regulation of feeding behavior and body temperature. The activity of GABAergic neurons of the arcuate nucleus was depressed upon feeding and was similarly depressed upon oral administration of MDP. Infusion of MDP in single neurons and patch-clamp recording of neuron excitability demonstrated that MDP-mediated control of GABAergic neurons was cell-autonomous.

We next tested whether the expression of Nod2 in hypothalamic GABAergic neurons was necessary to control food intake and body temperature. Indeed, the ablation of the *Nod2* gene in hypothalamic neurons alone using the local injection of a Cre-expressing virus led to weight gain in older mice. Moreover, this treatment altered nest building behavior and body temperature control. The intestinal microbiota is the most probable source of Nod2 ligand in this context, as oral antibiotic treatment abrogated the Nod2-mediated control over feeding.

CONCLUSION: Our study shows that structural components of the bacterial microbiota can be directly sensed by hypothalamic neurons to regulate feeding behavior, nesting behavior, and body temperature. In this way, intestinal bacteria may be used by the brain as an indirect measure of food intake or as a direct measure of bacterial expansion or death attributable to food intake. In the latter scenario, bacterial expansion or death may be associated with perturbation of intestinal homeostasis or a risk of pathogenesis. Alternatively, resident bacteria may regulate food intake to protect their intestinal niche. ■



Metabolic control via the gut-brain axis. Food consumption induces expansion of the intestinal microbiota. This expansion is followed by an increase in muropeptide release from the gut bacteria. When they reach the brain, these muropeptides target a subset of inhibitory hypothalamic neurons. In older females, activation of neuronal Nod2 receptors by muropeptides decreases neuronal activity, which in turn helps to regulate satiety and body temperature.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: ilanaga@gmail.com (I.G.); gerard.eberl@pasteur.fr (G.E.); pierre-marie.lledo@pasteur.fr (P.-M.L.)
†These authors contributed equally to this work.
Cite this article as I. Gabanyi *et al.*, *Science* **376**, eabj3986 (2022). DOI: 10.1126/science.abj3986

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abj3986

RESEARCH ARTICLE

NEUROIMMUNOLOGY

Bacterial sensing via neuronal Nod2 regulates appetite and body temperature

Ilana Gabanyi[1,2]*, Gabriel Lepousez[1], Richard Wheeler[3], Alba Vieites-Prado[4], Antoine Nissant[1], Sébastien Wagner[1], Carine Moigneu[1], Sophie Dulauroy[2], Samia Hicham[3], Bernadette Polomack[2], Florine Verny[4], Philip Rosenstiel[5], Nicolas Renier[4], Ivo Gomperts Boneca[3], Gérard Eberl[2]*†, Pierre-Marie Lledo[1]*†

Gut bacteria influence brain functions and metabolism. We investigated whether this influence can be mediated by direct sensing of bacterial cell wall components by brain neurons. In mice, we found that bacterial peptidoglycan plays a major role in mediating gut-brain communication via the Nod2 receptor. Peptidoglycan-derived muropeptides reach the brain and alter the activity of a subset of brain neurons that express Nod2. Activation of Nod2 in hypothalamic inhibitory neurons is essential for proper appetite and body temperature control, primarily in females. This study identifies a microbe-sensing mechanism that regulates feeding behavior and host metabolism.

Brain homeostasis and its downstream effects are sensitive to gut microbiota (*1*). In the absence of microbiota, brain chemistry and metabolism are altered, leading to cognitive and behavioral dysfunction (*2*, *3*). Secreted bacterial compounds found in the circulation have been implicated in microbiota-brain communication pathways and have been used or targeted to treat brain-related disorders (*4*–*6*). During homeostasis, the composition of the gut microbiota changes constantly (*7*), leading to the cyclic release of bacterial compounds into the gut lumen. Some of these compounds can influence metabolism, the immune system, and behavior in humans and mice (*2*). One such compound, peptidoglycan (PG), is a major component of the bacteria cell wall. Fragments of PG are released upon bacterial growth, replication, or death (*8*). PG fragments known as muropeptides have been found in mouse brains (*9*), and studies in *Drosophila* have demonstrated their capacity to influence neuronal activity and plasticity (*10*). Because they are present in almost all bacteria and are constantly released, muropeptides may serve as important gut-derived signals to the brain.

In mammals, muropeptides are recognized by cytosolic Nod-like receptors (Nod1 and Nod2) and by peptidoglycan recognition proteins (PGRPs) (*11*). Nod1 recognizes *meso*-diaminopimelic acid (*meso*-DAP)–containing muramyl tripeptides derived mainly from Gram-negative bacteria, whereas Nod2 recognizes muramyl dipeptide (MDP), a motif found in every bacterial PG type (*12*).

Nod2 and its ligands are associated with neurodegeneration and memory functions in mouse models of Parkinson's (*13*) and Alzheimer's diseases (*14*). In humans, variants of *NOD2* are associated with bipolar disorder, schizophrenia, and Parkinson's disease (*15*–*17*). In addition, muropeptides have been implicated in sleep alteration (*18*), and Nod2 and MDP contribute to metabolic regulation (*19*, *20*). Nod2 deficiency leads to metabolic dysfunction in response to diet-induced obesity (*19*), whereas MDP exhibits a protective role in obesity-induced insulin resistance (*20*). Thus, Nod2 signaling plays a role in both brain and metabolic pathologies. However, it is unknown whether a gut-brain pathway involving neuronal responses to Nod2 activation is necessary to maintain physiological homeostasis.

Here, we identified a neuronal and cell-autonomous Nod2 signaling pathway and explored the impact of this pathway on brain activity and its consequences on behavior and metabolism. We show that Nod2 is expressed by a subset of hypothalamic neurons that respond to MDP from the intestine and regulate food consumption, body temperature, and associated behaviors. This work uncovers a bacteria-driven gut-brain communication modality involved in the control of energy homeostasis.

Brain neurons express Nod2

We first investigated the expression pattern of Nod2 in the central nervous system (CNS) using heterozygote knock-in mice that harbor one allele encoding a functional Nod2 receptor and the other allele encoding green fluorescent protein (GFP) (*Nod2*^{tm1Jhgt}, hereafter *Nod2*^{GFP} mice), in which GFP serves as a reporter for Nod2 expression. We detected GFP expression in several brain regions and by distinct cell types (Fig. 1, A to D, and fig. S1, A to C). Neurons expressing GFP, variable in morphology and size, were found mostly in the striatum, thalamus, and hypothalamus (Fig. 1, B to D, and fig. S1B). No significant neuronal expression was found in the cortex (Fig. 1, A to D, and fig. S1A). In contrast to this selective neuronal expression, microglial and endothelial cells expressing GFP were found in all brain regions (Fig. 1, C and D, and fig. S1C). The neuronal expression of GFP did not extend to the intestine, where strong GFP expression was detected in endothelial cells (fig. S1D). This pattern of *Nod2* expression in specific brain regions was confirmed using *Nod2* mRNA in situ hybridization (fig. S2).

Microbiota-derived muropeptides are found in the brain

To determine whether Nod2 ligands from the intestine could directly regulate brain neuronal activity, we first assessed whether orally administered muropeptides reach the brain. To this end, mice were gavaged with radiolabeled muropeptides, and tissues were collected 4 hours later (fig. S3A). Muropeptides were able to cross the gut barrier, reach the blood circulation, and accumulate in the brain (fig. S3, B and C). Female mice accumulated more muropeptides in the blood than males, but no differences were detected in the brain (fig. S3, B and C). To assess trafficking to the brain of muropeptides released by gut-resident bacteria, we colonized mice with *Escherichia coli* containing radiolabeled PG and then examined tissues after 24 hours (Fig. 1E). Females accumulated more muropeptides in the brain than males, even though similar amounts of muropeptides were detected in the blood in both groups (Fig. 1F and fig. S3, D to G). Thus, muropeptides can reach the brain from the gut, possibly at different rates in males and females.

Lack of Nod2 on GABAergic neurons leads to metabolic alterations

We next assessed whether a loss in neuronal expression of Nod2 affects brain-controlled metabolism and behavior. Mice were generated that lacked expression of Nod2 in two main classes of CNS neurons: inhibitory vesicular γ-aminobutyric acid transporter–positive (Vgat^{+}) neurons (henceforth GABAergic neurons) and excitatory calcium/calmodulin-dependent protein kinase II (CamKII^{+}) neurons. To this end, mice encoding floxed alleles of *Nod2* (*Nod2*^{flox} mice) were crossed to *Vgat*^{cre} (*Slc32a1*$^{tm2(cre)Lowl}$) or *Camk2a*^{cre} mice to generate *Vgat*$^{\Delta Nod2}$ and *CamKII*$^{\Delta Nod2}$ mice. Over a

[1]Institut Pasteur, Université Paris Cité, CNRS UMR 3571, Perception and Memory Unit, F-75015 Paris, France. [2]Institut Pasteur, Université Paris Cité, INSERM U1224, Microenvironment and Immunity Unit, F-75015 Paris, France. [3]Institut Pasteur, Université Paris Cité, CNRS UMR6047, INSERM U1306, Biology and Genetics of the Bacterial Cell Wall Unit, F-75015 Paris, France. [4]Sorbonne Université, Paris Brain Institute–ICM, INSERM U1127, CNRS UMR7225, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013 Paris, France. [5]Institute of Clinical Molecular Biology, Christian-Albrechts-Universität zu Kiel and University Hospital Schleswig-Holstein, Campus Kiel, 24105 Kiel, Germany.
*Corresponding author. Email: ilanaga@gmail.com (I.G.); gerard.eberl@pasteur.fr (G.E.); pierre-marie.lledo@pasteur.fr (P.-M.L.)
†These authors contributed equally to this work.

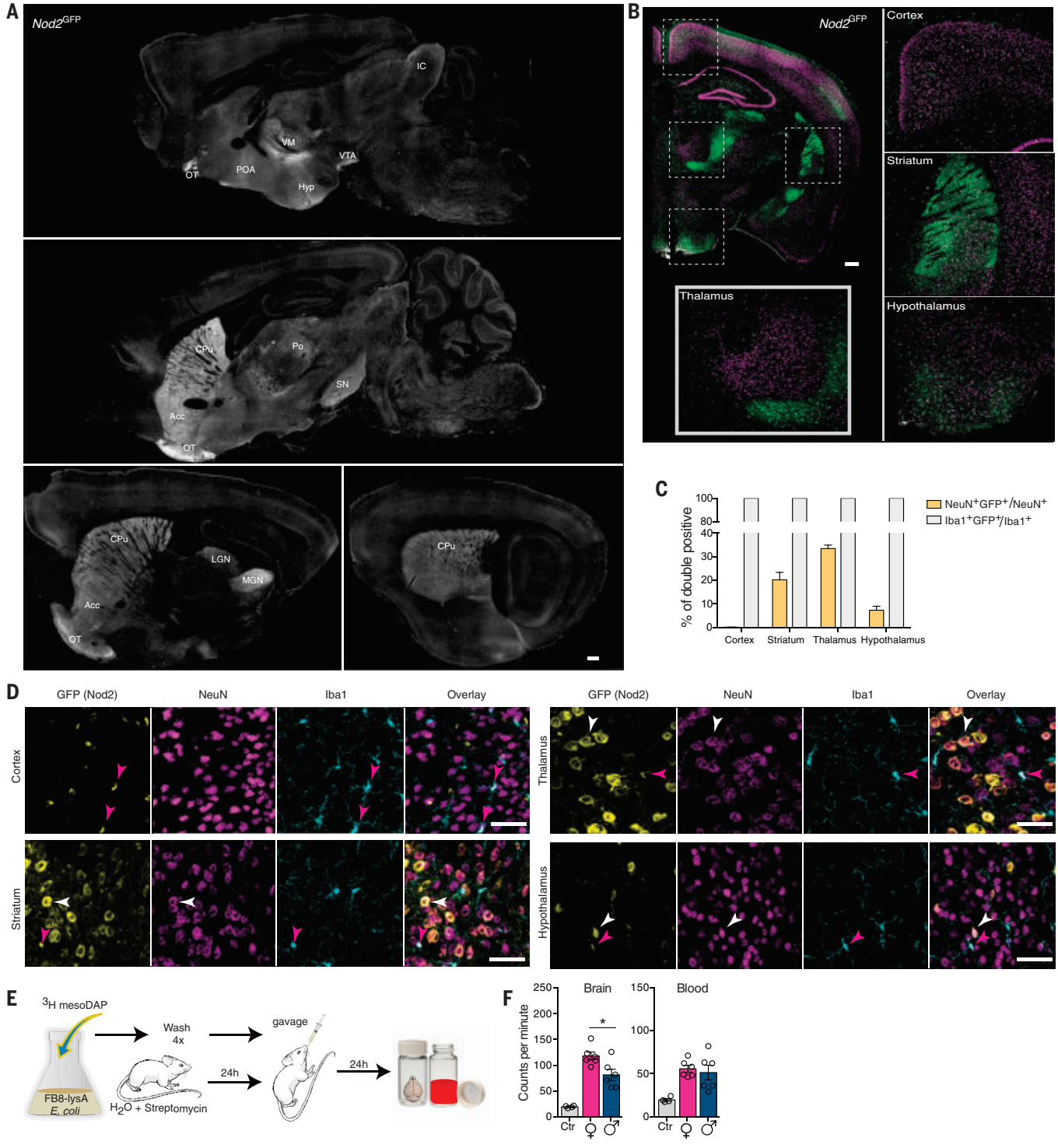


Fig. 1. Nod2 and Nod2 ligands in the brain. (**A** and **B**) Immunofluorescence of *Nod2*^{GFP} mouse brain slices, as shown in representative images of brain sagittal slices stained for Nod2 (GFP, white) (A) and a coronal slice stained for neurons (NeuN, magenta), microglia (Iba1, white), and Nod2 (GFP, green) (B), highlighting cortex, thalamus, striatum, and hypothalamus. Scale bars, 150 μm. (**C**) Quantification of neurons (NeuN $^{+}$), microglia (Iba1^{+}), and their colocalization with Nod2 (GFP^{+}) in the cortex, striatum, thalamus, and hypothalamus. (**D**) Colocalization of Nod2 (yellow) with Iba1 (cyan) or NeuN (magenta) in the cortex, thalamus, striatum, and hypothalamus. Pink arrowheads indicate Iba1^{+}Nod2^{+} cells; white arrowheads indicate NeuN^{+}Nod2^{+} cells. Scale bars, 10 μm. (**E**) PG tracking experiment scheme. (**F**) Radioactivity detected 24 hours after gavage with ^{3}H-labeled *E. coli* in the brain and blood of older females and males (7 to 8 months; n = 6 per group). Controls were gavaged with nonlabeled *E. coli* (n = 4). Data are averages ± SEM. *$P \leq 0.05$ (unpaired t test). Abbreviations: CPu, caudate putamen; Hyp, hypothalamus; IC, inferior colliculus; LGN, lateral geniculate nucleus; MGN, medial geniculate nucleus; Acc, nucleus accumbens; OT, olfactory tubercles; Po, posterior thalamic nucleus; POA, preoptic area; SN, substantia nigra; VM, ventral medial nucleus of the thalamus; VTA, ventral tegmental area; PG, peptidoglycan.

period of several months, *Vgat*$^{\Delta Nod2}$ females gained more weight than did *Nod2*^{flox} and *Vgat*^{cre} controls (Fig. 2A). Weight differences became significant at ~6 months of age and further increased as mice aged (Fig. 2A). No weight differences were observed in *Vgat*$^{\Delta Nod2}$ males or in *CamKII*$^{\Delta Nod2}$ mice of either sex (Fig. 2, A and B).

In agreement with these findings, an increase in appetite was observed only among older (>6 months) *Vgat*$^{\Delta Nod2}$ female mice (Fig. 2C and fig. S4, A and B). *Vgat*$^{\Delta Nod2}$ mice ate a greater number of food pellets relative to *Nod2*^{flox} mice (Fig. 2D). Although control mice ate more frequently, the number of pellets eaten during each meal bout was significantly higher in *Vgat*$^{\Delta Nod2}$ female mice (Fig. 2E). Thus, the absence of Nod2 in inhibitory neurons of older female mice leads to a delay in achieving satiety, revealing a role for Nod2^{+} GABAergic neurons in appetite control. In mice that fully lacked Nod2 expression (*Nod2*^{KO}), both females and males showed increased weight gain with age (fig. S4, C and D); this finding suggests that additional effects on eating behavior and weight gain are regulated by non-neuronal and intestinal expression of Nod2 (*21*).

To assess whether the Nod2 ligand MDP modulates appetite control, we gavaged *Vgat*$^{\Delta Nod2}$ and *Nod2*^{flox} littermate control mice with MDP (L-D isoform) or an MDP isomer (L-L isoform, hereafter termed MDPctr) that does not activate Nod2 (*12*). We observed that control mice treated with MDP ate less than when treated with MDPctr, whereas *Vgat*$^{\Delta Nod2}$ mice ate more (Fig. 2F). Thus, MDP acts as a satiety signal in mice via the Nod2 receptor expressed by inhibitory neurons.

Older *Vgat*$^{\Delta Nod2}$ females also displayed alterations in nest-building behavior, shredding fewer cotton nestlets than *Nod2*^{flox} and *Vgat*^{cre} mice (Fig. 2G and fig. S4E). In rodents, nesting behavior is involved in heat conservation and is strongly influenced by environmental and body temperature (*22*). In agreement with the altered nesting behavior, the daily temperature variation was smaller in *Vgat*$^{\Delta Nod2}$ females

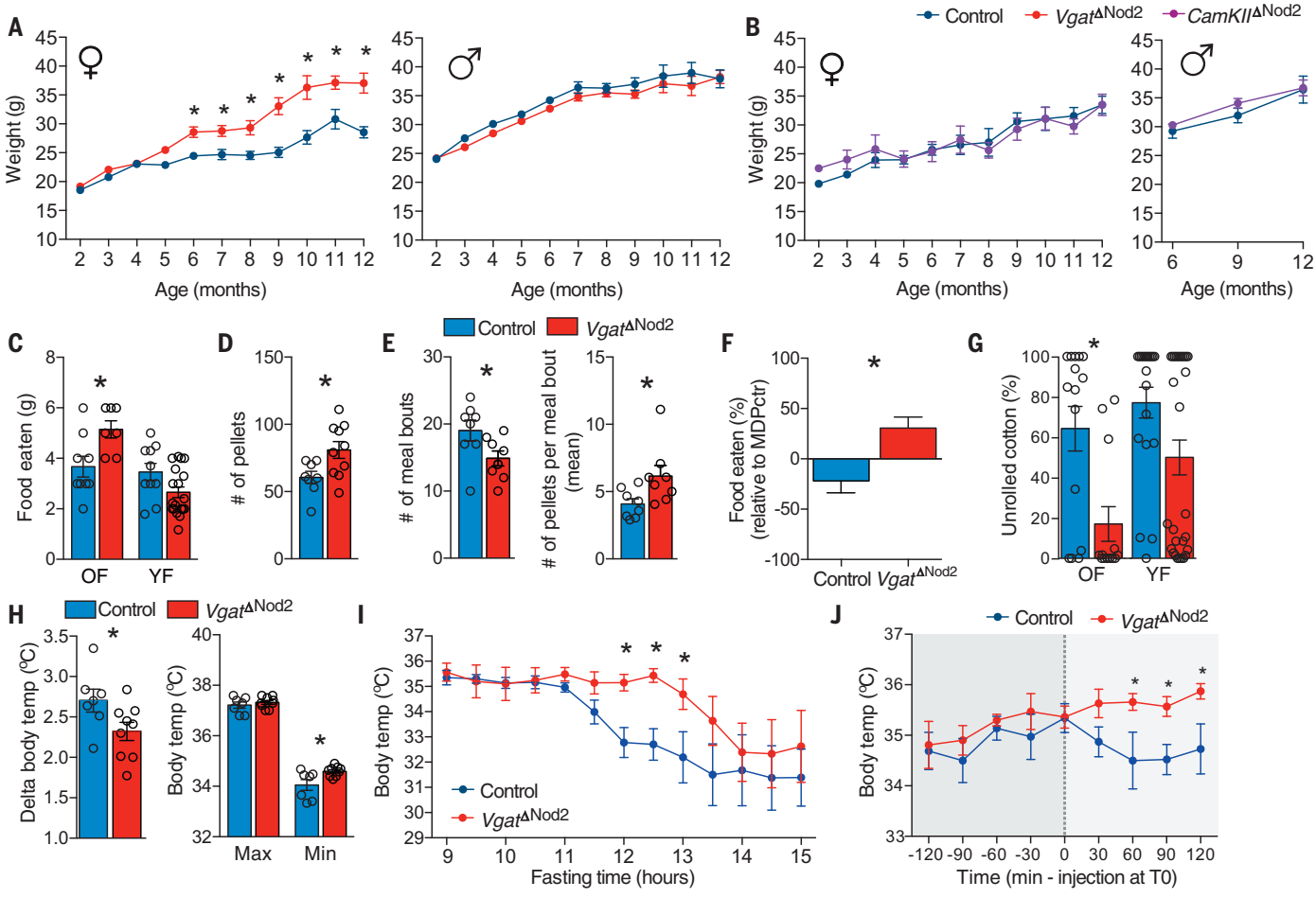


Fig. 2. Nod2 expression by inhibitory neurons affects feeding and temperature in older female mice. (**A** and **B**) Mouse body weights up to 12 months of age. (A) *Vgat*$^{\Delta Nod2}$ females (n = 13 to 30) and males (n = 6 to 29). (B) *CamKII*$^{\Delta Nod2}$ females (n = 3 to 14) and males (n = 7 to 9). (**C** to **F**) Eating behavior of *Vgat*$^{\Delta Nod2}$ female mice. (C) Food eaten in 24 hours by older (7 to 8 months; n = 7 to 9) and younger mice (3 to 4 months; n = 10 to 19). [(D) to (F)] Food eaten measured with an automated system during the dark period (age >6 months). (D) Number of food pellets eaten. (E) Meal bouts (left) and food pellets eaten per meal (right; n = 8 per group). (F) Food pellets eaten 4 hours after MDP or MDPctr gavage (n = 5 or 6). Data were normalized by amount eaten after MDPctr gavage. (**G**) Nest-building test. Amount of cotton used to build the nest (unrolled cotton) is shown for younger (3 to 4 months; n = 20 to 29) and older mice (7 to 8 months; n = 13 to 15). (**H** to **J**) Body temperature measurements using telemetric probes for *Vgat*$^{\Delta Nod2}$ females (age >6 months). (H) Daily variation in temperature at steady state. Left: Daily difference between maximum and minimum temperature (delta) per mouse. Right: Mean maximum and minimum reached over several days (n = 7 to 9). (I) Body temperature response to fasting (n = 4 to 6). (J) Body temperature response before and after β_3-adrenergic agonist (CL316,243) injection (n = 5 per group). Data are averages ± SEM. *$P \leq 0.05$ [two-way ANOVA in (A), (B), (C), (I), and (J); unpaired t test in (D), (E), (F), and (H) for three variables, delta, max, and min; Fisher's exact test in (G)]. Abbreviations: OF, older females; YF, younger females; MDP, muramyl dipeptide; MDPctr, MDP isomer.

than in control mice (Fig. 2H and fig. S4F). Moreover, female $Vgat^{\Delta Nod2}$ mice showed delayed temperature drops in response to fasting (Fig. 2I and fig. S4G). Similarly, β_3-adrenergic agonist injection induced rapid temperature drops in control but not in $Vgat^{\Delta Nod2}$ mice (Fig. 2J). Together, these results show that $Vgat^{\Delta Nod2}$ female mice developed altered body temperature control. Older $Vgat^{\Delta Nod2}$ female mice eventually developed a diabetic-like phenotype (fig. S4H) and exhibited reduced lifespans (fig. S4I). Thus, Nod2 expression by inhibitory neurons plays an important role in the control of female metabolism.

Sex- and age-dependent MDP-mediated activation of brain neurons

To identify the brain regions affected by MDP, as well as to understand why MDP regulated feeding and body temperature only in older females, younger (2 to 3 months) females and older (7 to 8 months) females and males were gavaged with either the Nod2 ligand MDP or MDPctr. An unbiased mapping of Fos expression in the brain revealed that MDP administration induced distinct patterns of neuronal activation among younger and older females and males, as well as a more pronounced effect in older mice, where numerous nuclei were affected (Fig. 3A and table S1). These age-related differences were not associated with a failure of gut-derived muropeptides to reach the brain of younger females (fig. S5, A to F). After MDP gavage, only older females showed significant alterations in neuronal activity of the arcuate (ARC) and dorsomedial (DMH) nuclei of the hypothalamus as well as in the lateral hypothalamic area, key regions involved in the regulation of feeding behavior and body temperature (Fig. 3, A and B). Thus, older females exhibit higher responsiveness to MDP in regions involved in the regulation of appetite and body temperature (*23*, *24*). Finally, we did not observe significant differences with age and/or sex for neuronal Nod2 expression in the ARC or DMH (fig. S5, G and H).

Hypothalamic GABAergic neurons respond to MDP

To further dissect the MDP-mediated effects on hypothalamic neurons, we analyzed the effect of MDP on GABAergic (inhibitory) neurons of the ARC ($Vgat^{ARC}$). We first confirmed by in situ hybridization that $Vgat^+$ and neuropeptide Y (NPY^+) neurons in this area expressed Nod2 (Fig. 4A and fig. S6A). By contrast, pro-opiomelanocortin (POMC)–expressing (non-GABAergic) neurons did not express Nod2 (fig. S6B). The specific expression of the calcium sensor GCaMP in GABAergic neurons allowed for the monitoring of $Vgat^{ARC}$ neuronal activity (Fig. 4, B and C). These neurons are active after a fasting period and rapidly decrease activity upon feeding (*25*) (Fig. 4, D and E). MDP but not MDPctr administration induced a similar long-lasting decrease in the spontaneous activity of $Vgat^{ARC}$ neurons in control mice, but not in $Vgat^{\Delta Nod2}$ mice (Fig. 4, D to G, and fig. S6, C to I). Thus, MDP decreases $Vgat^{ARC}$ neuronal activity via the Nod2 receptor and influences appetite control by modulating hypothalamic circuits.

To demonstrate cell-autonomous regulation of $Vgat^{ARC}$ neurons by MDP, we performed ex vivo patch-clamp recordings on brain slices from $Vgat^{cre}Nod2^{GFP}$ mice injected with a Cre-dependent reporter virus to target $Vgat^+Nod2^+$ neurons and confirm their identity after recording (fig. S6J). We characterized cell excitability after infusion of MDP or MDPctr into individual neurons. Thirty minutes after infusion, MDP, but not MDPctr, induced a strong decrease in the number of action potentials elicited in Nod2-expressing neurons (Fig. 4, H and I, and fig. S6K). This reduction in cell excitability did not result from changes in cell membrane–intrinsic properties, as no significant differences were observed in input membrane resistance (Fig. 4J), nor in the threshold to trigger firing activity (Fig. 4K). Thus, MDP decreases the $Vgat^{ARC}$ neuronal activity in a cell-autonomous manner.

Expression of Nod2 in ARC and DMH neurons regulates body weight and temperature

Finally, we assessed whether Nod2-expressing hypothalamic neurons regulated metabolism in "steady state" control female mice. In such mice, the ARC and DMH neurons showed different levels of activity, as measured by Fos expression, relative to $Vgat^{\Delta Nod2}$ mice (Fig. 5, A and B, and table S2). We observed in the ARC and DMH nuclei $Nod2^+$ and $Nod2^+Vgat^+$ neurons (Fig. 5, C to F). To confirm the causal implication of these neurons in the control of metabolism in older females, we locally abrogated Nod2 expression by the injection of a Cre-expressing virus into the hypothalamus of $Nod2^{flox}$ mice (Fig. 5, G and H). Cre-injected $Nod2^{flox}$ females gained more weight, consumed more food, and showed a smaller variation in body temperature than Cre-injected controls (Fig. 5, I to K). These mice also developed a tendency to decreased nest building (Fig. 5L). Thus, the expression of Nod2 in ARC and DMH hypothalamic neurons is necessary to regulate feeding behavior and body temperature in older female mice.

To confirm that microbiota-derived Nod2-ligands are involved in this regulation, we injected $Nod2^{flox}$ and control female mice with Cre-expressing virus and treated them with broad-spectrum antibiotics (ABX) for 13 weeks. ABX treatment has been shown to effectively reduce the amount of muropeptides in the blood (*26*). No difference in weight gain or food consumption was observed between $Nod2^{flox}$ and control mice injected with the Cre-expressing virus during ABX treatment (Fig. 5, M and N, and fig. S7A). By contrast, ABX removal led to increased weight gain in $Nod2^{flox}$ mice (Fig. 5M) and to a decrease in food consumption only in control mice (Fig. 5N), which was associated with a normalization of the intestinal microbiota (fig. S7, B and C). Thus, the microbiota plays a role in the production of Nod2 ligands and the regulation of appetite by Nod2-expressing hypothalamic neurons.

Discussion

This work reveals a gut-brain communication pathway in which the expression of the Nod2 receptor in hypothalamic inhibitory neurons regulates appetite and body temperature in response to bacterial-derived muropeptides.

Using mutant mice and virus-induced gene deletion, we have identified the role of Nod2 in inhibitory neurons in the control of body temperature and appetite. Diverse mechanisms have been proposed for the bacterial influence on host appetite, involving microbial metabolites such as short-chain fatty acids (*27*) and the *E. coli* protein ClpB (*28*). Here, we describe another mechanism by which gut bacteria muropeptides control host feeding behavior. As the transient postprandial increase in the gut microbial population (*7*) may lead to an increased release in cell wall–derived muropeptides, the host may use this bacterial signal to limit feeding as well as bacterial expansion. Alternatively, the bacterial microbiota may modulate the host's feeding behavior to stabilize its intestinal niche.

Previous studies, using loss-of-function approaches, have also reported diverse functions of PG in gut-brain cross-talk. Lack of proper PG cleavage by the host (due to the absence of Pglyrp2) leads to several behavioral impairments including anxiety-like behavior (*29*). Moreover, the specific deletion of the Nod1 receptor in epithelial cells increases the susceptibility of mice to stress-induced anxiety-like behavior and cognitive impairment (*30*). In our hands, no changes in anxiety-like behavior were observed in $Vgat^{\Delta Nod2}$ mice or in $CamKII^{\Delta Nod2}$ female mice (fig. S8, A and B). By contrast, total $Nod2^{KO}$ females over 6 months of age developed stronger anxiety-like behavior relative to their wild-type controls (fig. S8C). Given the widespread expression of Nod2 in different cell types and locations, the differences observed between $Vgat^{\Delta Nod2}$ and $Nod2^{KO}$ mice suggest that additional PG-dependent alterations may arise from mechanisms involving non-neuronal cell types. In addition to bacterially derived muropeptides, endogenous Nod2 ligands may also play a role in the phenotypes we observed. Although its relevance in vivo or in neurons has not yet been demonstrated, in vitro studies have revealed that Nod2 may respond to endogenous

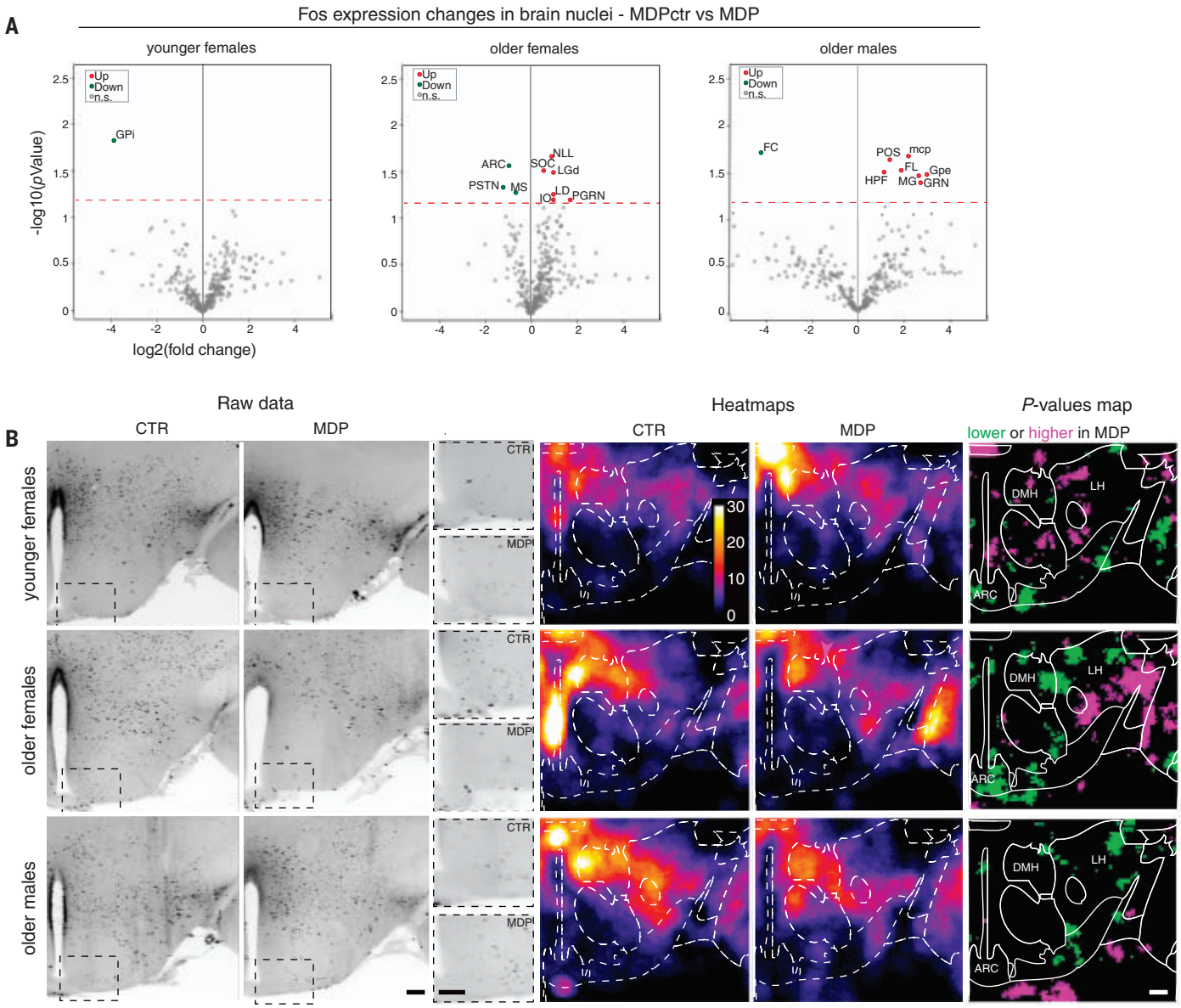


Fig. 3. MDP affects neuronal activation in a sex- and age-dependent manner. (**A** and **B**) Analysis of brain neuron Fos expression in wild-type mice 3 hours after MDP or MDPctr gavage of younger females (2 to 3 months), older females (7 to 8 months), and older males (7 to 8 months) (n = 4 per group). (A) Volcano plots from the automated analysis of Fos^+ cells distribution in the brain. (B) Raw data and heatmaps indicating the number of Fos^+ cells, and P-value maps of Fos expression in the hypothalamus, highlighting DMH, ARC, and LH. Shown are regions with higher (magenta) or lower (green) numbers of Fos^+ neurons in the MDP group as compared to the MDPctr group; scale bars, 200 μm. *$P \leq 0.05$ (unpaired t test).

Abbreviations: ARC, arcuate hypothalamic nucleus; DMH, dorsomedial nucleus of the hypothalamus; LGd, dorsal part of the lateral geniculate complex; FC, fasciola cinereal; FL, flocculus; GRN, gigantocellular reticular nucleus; Gpe, globus pallidus external segment; GPi, globus pallidus internal segment; HPF, hippocampal formation; IO, inferior olivary complex; LD, lateral dorsal nucleus of thalamus; LH, lateral hypothalamic area; MG, medial geniculate complex; MS, medial septal nucleus; mcp, middle cerebellar peduncle; NLL, nucleus of the lateral lemniscus; PGRN, para-gigantocellular reticular nucleus; PSTN, parasubthalamic nucleus; POS, postsubiculum; SOC, superior olivary complex; CTR, MDPctr.

sphingosine-1-phosphate released under cellular stress conditions (*31*).

The lack of Nod2 in inhibitory neurons had particularly strong effects on appetite in females over 6 months of age. A variety of brain- and metabolism-related diseases have shown sex- and age-dependent phenotypes (*32*–*35*). Thus, understanding the mechanisms behind such biases may lead to more specific and efficient therapeutic approaches. To explore the mechanisms behind such sex- and age-based specificity, we evaluated different parameters: (i) Nod2 expression by hypothalamic nuclei, (ii) brain neuronal responses to MDP, and (iii) PG pharmacokinetics. Among these factors, only the increased accumulation of muropeptides in the female brain (relative to males) correlates with sex-specific hypothalamic neuronal activation in response to MDP. Previous studies have also described age and sex biases in the context of gut-brain communication (*29*, *36*). For example, microglial responses to bacterial metabolites are stronger in males during the embryonic phase, whereas female microglia are more responsive in adulthood (*36*). Moreover, the absence of Pglyrp2 influences brain development and behavior, leading eventually to sex- and age-dependent alterations (*9*, *29*). Many additional factors that show age and sex differences—such as hormonal status and

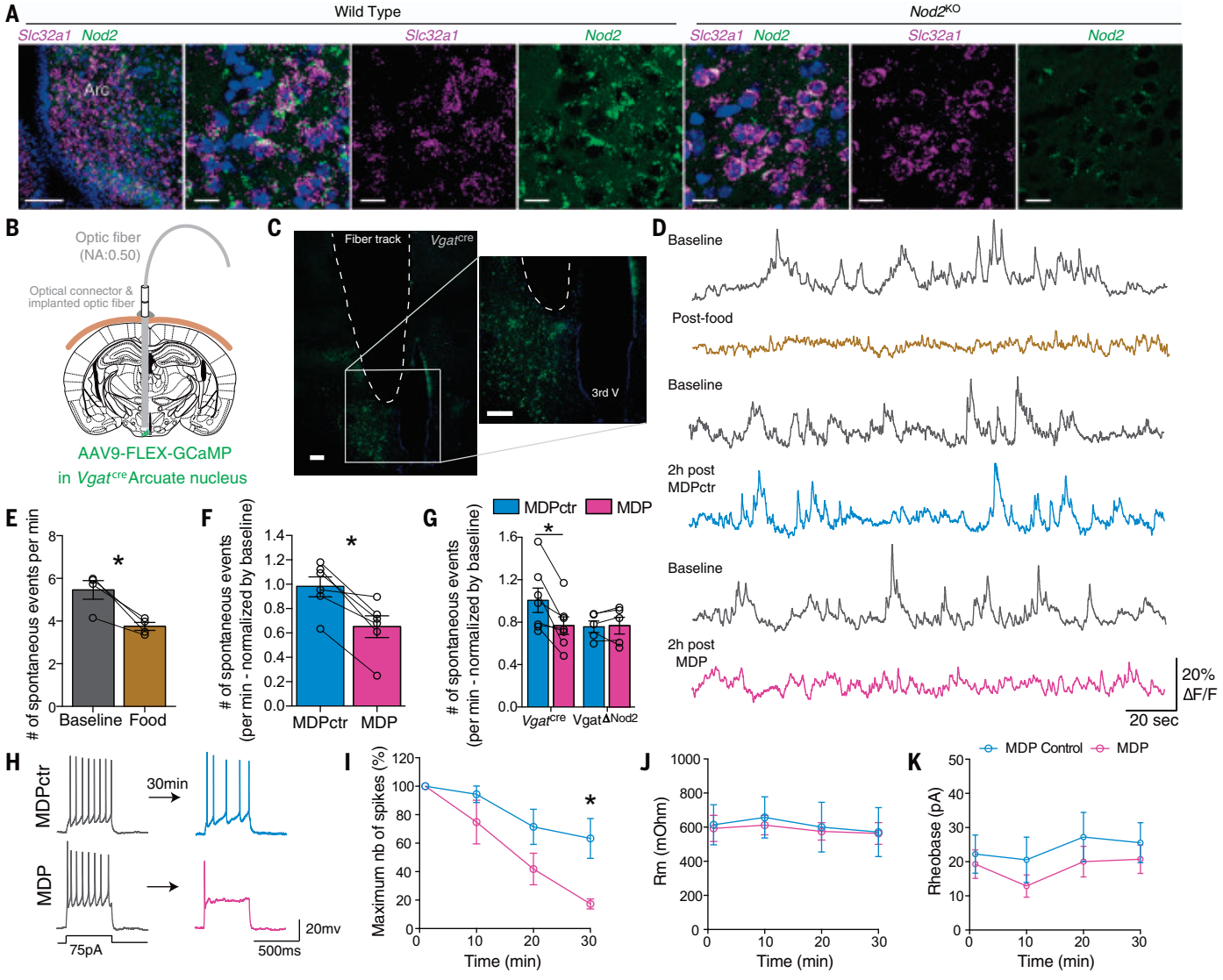


Fig. 4. MDP decreases Vgat^{ARC} neuronal activity. (**A**) Representative images of RNAscope with *Nod2* (green) and *Slc32a1* (*Vgat*, magenta) probes and nuclear staining (DAPI, blue) in the ARC. Colocalization of *Nod2* and *Slc32a1* can be seen, with *Nod2*^{KO} mice used as a negative control. Scale bars, 25 μm (except first image, 100 μm). (**B** to **G**) In vivo fiber photometry (females >6 months). (B) Scheme of calcium activity recording scheme for Vgat^{ARC} neurons. (C) Representative immunofluorescence image of the brain of a mouse subjected to fiber photometry, showing expression of GCaMP (GFP, green); cell nucleus is stained with DAPI (blue). Scale bars, 200 μm. Injection site and canula track (dashed lines) are highlighted; 3rd V, third ventricle. (D) Representative traces of GCaMP6f fluorescence signals from Vgat^{ARC} neurons of fasted mice. (E) Number of spontaneous events before (baseline) and after food presentation (food) for fasted mice. Lines connect each individual (n = 4). (F) Number of spontaneous events after MDPctr or MDP gavage for fasted *Vgat*^{cre} mice (n = 6). (G) Number of spontaneous events without fasting for *Vgat*^{cre} (n = 7) and *Vgat*$^{\Delta Nod2}$ (n = 5) mice. Lines connect each individual. Data were normalized by baseline obtained before gavage. (**H** to **K**) Ex vivo patch-clamp recordings in Nod2^{+} Vgat^{ARC} neurons (females >6 months; n = 4 mice; nine MDPctr and seven MDP cells). (H) Representative traces of triggered action potentials from cells treated with MDPctr or MDP at 0 and 30 min. (I) Maximum number of spikes, (J) membrane resistance, and (K) rheobase measurements along the recording. Data are averages ± SEM. *$P \leq 0.05$ [paired t test in (E) and (F); two-way ANOVA in (G), (I), (J), and (K)].

intestinal and blood-brain barrier permeability, as well as microbiota composition—are also likely to play a role (*32*, *37*–*39*).

From an evolutionary perspective, female-specific appetite control systems may have been positively selected because of the greater importance of energy balance in females for sexual maturity and pregnancy (*40*). In our study, the differences in weight became significant around 6 months of age and increased with time. This time period correlates with hormonal changes in female mice associated with pre-menopausal states in humans (*41*), such as the substantial decrease in estradiol production. This hormone plays an important role not only in sexual maturity but also in energy balance. Estradiol (E2) controls food intake via NPY and POMC neurons. Specific deletion of estrogen receptor α (ERα) in POMC neurons leads to body weight increase in female mice, and E2-induced anorexia is blunted in mice lacking NPY neurons (*42*). Our findings show that MDP's effects on food intake are mediated via the GABAergic neurons in the ARC, which includes NPY neurons. Therefore, estradiol and MDP may exert their effects through similar pathways. To explain the appearance of this phenotype in older females alone, we hypothesize that MDP's effect on food intake is masked by the presence of higher levels of estradiol in younger females, as estradiol may be a more potent inducer of anorexia. As the levels of estradiol begin to

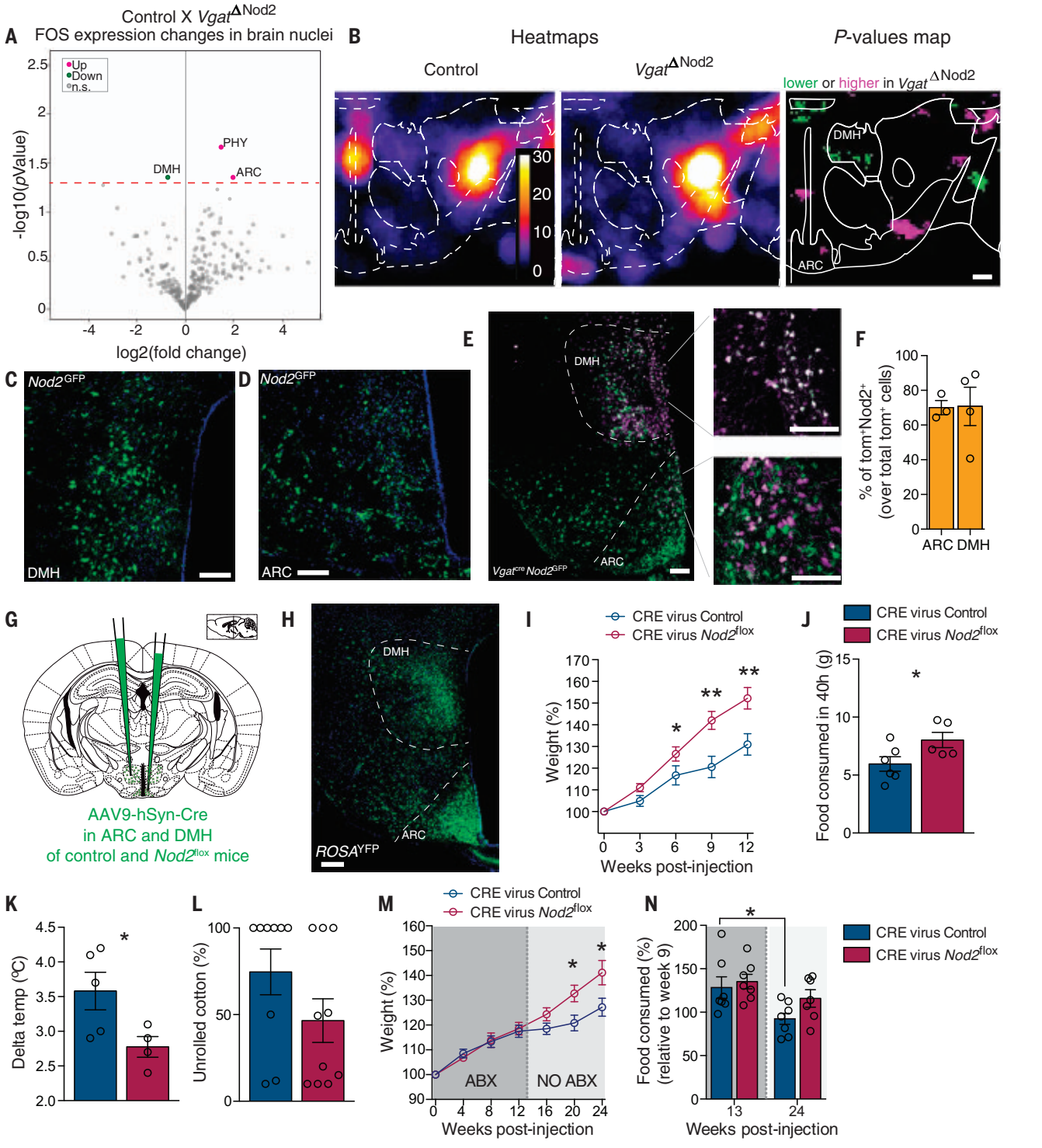


Fig. 5. Hypothalamic Nod2^{+} neurons are required for weight and temperature control. (**A** and **B**) Fos expression in female *Vgat*$^{\Delta Nod2}$ mice and controls at steady state (7 to 8 months; n = 4 per group). (A) Volcano plots of brain Fos^{+} cell distribution in the brain. (B) Heatmaps indicating the number of Fos^{+} cells and P-value map of Fos expression in the hypothalamus, highlighting DMH and ARC. Shown are regions with higher (magenta) or lower (green) numbers of Fos^{+} neurons in the *Vgat*$^{\Delta Nod2}$ group as compared to the control group. Scale bar, 200 μm. (**C** and **D**) Representative immunofluorescence images of ARC (C) and DMH (D) of *Nod2*^{GFP} mouse, showing expression of Nod2 (GFP, green); cell nucleus is stained with DAPI (blue). (**E**) Representative immunofluorescence image of *Vgat*^{cre}*Nod2*^{GFP} mouse brain injected with AAV9.CAG.FLEX.Tdtomato virus in DMH and ARC. Colocalization of Nod2 (green) with Tdtomato (magenta) is highlighted; scale bars, 100 μm. (**F**) Frequency of Tdtomato^{+} cells expressing GFP among total Tdtomato^{+} cells (n = 4 mice). (**G**) Viral injection scheme. (**H**) Representative immunofluorescence image of *Rosa*^{YFP} mouse brain injected with AAV9.hSyn.Cre virus in DMH and ARC, showing expression of YFP (green); cell nucleus is stained with DAPI (blue). (**I**) Body weight weekly measures after viral injection (females, 2 to 3 months at week 0; n = 8 to 11). (**J**) Food eaten in 40 hours (n = 5 or 6). (**K**) Body temperature variation in

24 hours (n = 4 or 5). (**L**) Nest-building test. Amount of cotton used to build the nest (unrolled cotton) is shown (n = 9 or 10). (**M** and **N**) Virus-injected mice treated with ABX. (M) Body weight weekly measures during and after ABX treatment (females, 2 to 4 months at week 0; n = 13 per group). (N) Food eaten in 40 hours at week 13 (during ABX treatment) and week 24 (after ABX treatment) (n = 7 per group). Data are averages ± SEM. *$P \leq 0.05$ [unpaired t test in (A), (J), and (K); two-way ANOVA in (I), (M), and (N); Fischer's exact test in (L)]. Abbreviations: ARC, arcuate nucleus of the hypothalamus; DMH, dorsomedial nucleus of the hypothalamus; PHY, perihypoglossal nuclei; ABX, antibiotics.

decrease (as observed in pre-menopausal and menopausal states), this control of food intake would be attenuated and MDP's anorexic effects via Nod2 hypothalamic neurons would play a more prominent role in appetite control.

In addition to the well-studied control of host immune response by Nod2, our work highlights the need to consider the effects of Nod2 activation on brain neuronal activity in the context of using Nod2 ligands as a potential therapeutic tool. As PG-derived metabolites are already used in clinics for cancer therapies and planned for other applications (*43*), a better understanding of the physiological roles of Nod2 and its ligands is therefore extremely important. Our work reveals a sex- and age-dependent pathway of gut-brain cross-talk, which may open additional avenues for the treatment of neurological and metabolic disorders.

Materials and methods

Animals

Nod2$^{flox/flox}$ (*44*), *Nod2*^{tm1Jhgt} (*Nod2*^{GFP}) (*45*), and *CamK2A-iCRE* (*46*) mice were generously provided by P. Rosenstiel (University of Kiel), J. P. Hugot (Université Paris Diderot), and R. Liblau (Université Toulouse III), respectively. *Nod2*^{tm1Flv} (*Nod2*^{KO}), *Gt(ROSA)26Sor*$^{tm1(EYFP)Cos}$ (*ROSA*^{YFP}), and *Slc32a1*$^{tm2(cre)Lowl}$ (*Vgat*^{cre}) mice were purchased from Jackson Laboratories. These lines were interbred and/or backcrossed with C57Bl/6JRj at Institut Pasteur animal facilities to obtain the final strains described. Mice were maintained at Institut Pasteur animal facilities under specific pathogen-free conditions, with lights on at 7 a.m. and off at 9 p.m. Ages and sexes of mice are specified in the text or figure legends. Animal care and experimentation were approved by the committee on animal experimentation of Institut Pasteur (project DAP200025) and by the French Ministry of Research (MESR project 26737).

Immunofluorescence

To prepare brains for immunofluorescence imaging, mice were deeply anesthetized and intracardially perfused with 1× PBS for 5 min, then with 4% paraformaldehyde (PFA) in 0.1 M phosphate buffer for 10 min. The brains were removed and stored in PFA at 4°C until the following day. They were then washed three times with PBS and cryoprotected in 30% sucrose for 48 hours. Brain sections 60 μm thick were cut on a microtome. For gut immunofluorescence imaging, mice were killed by cervical dislocation and the small intestine was removed and placed in HBSS Mg2+Ca2+ (Gibco) + 5% FCS (complete media). The intestine was incised longitudinally and the luminal contents were washed away using complete media. The muscularis externa was then carefully removed from the underlying mucosa. The muscularis and mucosal tissue were pinned down on a plate coated with silicone (Smooth-ON; 83750B) and then fixed for 15 to 30 min with 4% PFA. Whole-mount samples were then permeabilized in 0.1% Triton X-100 for 1 to 2 hours. Brain slices and whole-mount intestinal samples followed the same staining procedure. Briefly, the tissue was rinsed, incubated in blocking buffer 10% normal serum + 0.1% Triton X-100 for 2 hours. It was then incubated with primary antibodies diluted in 1% normal serum + 0.1% Triton X-100 at appropriate concentrations and incubated overnight at 4°C. The following day the tissue was washed three times in 1× PBS and then incubated in 1% normal serum + 0.1% Triton X-100 with secondary antibodies. Samples were again washed once in 1× PBS, once in 1× PBS + DAPI (1:5000), and then once again in 1× PBS. Finally, tissue was mounted on slides with Fluoromount-G, cover slip added and sealed. Images were captured with a confocal laser-scanning microscope (LSM 700, Zeiss) using the 10×, 25×, or 40× Zeiss objectives [Plan-Apochromat 10×/0.3; I LCI Plan-Neofluar 25×/0.8 Imm Korr DIC M27; LD Plan-Neofluar 40×/1.3 Oil DIC (UV) VIS-IR M27], or a microscope equipped with Apotome system (Zeiss), a digital camera (Hamamatsu ORCA-ER C4742-80), and Axiovision 4.8 software (Zeiss), using the 10× and 20× Zeiss objectives (Fluar 10×/0.5 M27; Plan-Apochromat 20×/0.8 M27). Images were adjusted post hoc in ImageJ. The following primary antibodies were used: anti-GFP (chicken; 1:500; Millipore 06-896); anti-NeuN (mouse; 1:100; Millipore MAB377); anti-Iba1 (rabbit; 1:500; Wako 019-19741); anti-CD31 (rat; 1:50; BD Pharmaceutical 550274); anti-βIII Tubulin (rabbit; 1:1000; Cell signaling D71G9); anti-AgRP (goat; 1:2000; Neuromics GT15023); and anti-POMC (rabbit; 1:1,000, Phoenix Pharmaceuticals H02930). The following secondary antibodies from Molecular Probes were used at 1:1000: Alexa Fluor 488–conjugated goat anti-chicken IgG (A11039); Alexa Fluor 568-conjugated goat anti-mouse IgG (A21134); Alexa Fluor 568 or 647–conjugated goat anti-rabbit IgG (A11036 or A21244); and Alexa Fluor 647–conjugated goat anti-rat IgG (A21449).

Cell counting

Brain images were captured using a microscope equipped with Apotome system (Zeiss), a digital camera (Hamamatsu ORCA-ER C4742-80) and Axiovision 4.8 software (Zeiss), using the 10× (Fluar 10×/0.5 M27) or 20× (Plan-Apochromat 20×/0.8 M27) Zeiss objectives. Cell counting was performed manually using Fiji or the Icy open-source platform to draw regions and mark counted cells (http://icy.bioimageanalysis.org). To count GFP$^+$ cells in the cortex, striatum, thalamus, and hypothalamus, three female mice brains were stained with GFP, NeuN, and Iba1. For each animal, three to six random spots were selected and a minimum of 100 NeuN$^+$ and 100 Iba1$^+$ cells were then counted. Of those, the number of GFP$^+$ cells was counted. Neuron size was obtained by measuring a line drawn between the farthest points in the neuronal cell body using Fiji (9 to 12 neurons per animal). To count the GFP$^+$ neurons in the ARC and DMH, 12 brain sections from each *Nod2*^{GFP} mouse were selected. The slices were stained with GFP, NeuN, and AgRP (the latter being used to help determine the area of interest). The sex and age of the mice were revealed only after quantification. Data are shown as means per animal.

RNAscope in situ hybridization

RNAscope in situ hybridization was performed using mouse probes (Bio-Techne) against *Nod2* (*Mm-Nod2*, #433391), *Vgat* (*Mm-Slc32a1*-C2, #319191-C2), and *Npy* (*Mm-Npy-C2*, #313321-C2) on 15-μm sections of brain tissue according to the manufacturer's instructions.

Radiolabeling and purification of Lactobacillus rhamnosus strain Lr32 peptidoglycan

L. rhamnosus strain Lr32 was inoculated into 500 ml of 25% strength BD Difco MRS broth supplemented with 100 μM GlcNAc (Sigma-Aldrich). For radiolabeling, 10 μCi per liter of [^{14}C]*N*-acetylglucosamine (^{14}C-GlcNAc) was added. A nonlabeled control culture of Lr32 was prepared in parallel. Cultures were incubated overnight at 37°C without aeration and harvested at an OD_{600} of 2.5. Peptidoglycan (PG) was purified according to standard protocols for Gram-positive bacteria (*47*). Purified PG was stored short-term at −20°C. Approximately 40 mg of Lr32 PG was solubilized by resuspension with 12.5 mM NaH_2PO_4 pH 5.6 and incubated overnight at 37°C with 50 U of *Streptomyces globisporus* ATCC 21553 mutanolysin (Merck) per milligram of PG. Mice

were gavaged with 60,000 cpm of ^{14}C-GlcNAc-labeled PG. Control mice were gavaged with the equivalent dilution of nonlabeled PG. Collection of blood and organs was performed 4 hours after gavage, as described below.

Radiolabeling of PG in live streptomycin-resistant Escherichia coli *strain FB8-LysA*

The *E. coli* FB8-LysA strain is unable to convert *meso*-DAP to lysine. Thus, radioactive *meso*-DAP added to the growth media is incorporated specifically in the PG layer without the need for purification (*48*, *49*). A streptomycin-resistant strain of *E. coli* FB8-LysA was generated by plating the bacteria on LB agar containing a streptomycin sulfate gradient (0 to 50 mg/ml; Sigma-Aldrich) followed by incubation at 37°C. The colony exhibiting highest resistance to streptomycin after 24 hours was selected and amplified. Subculture of the resulting *E. coli* FB8-LysA Str^r strain confirmed that rapid growth was achieved in LB media containing streptomycin sulfate (50 mg/ml) and kanamycin (25 mg/ml; Sigma-Aldrich). For radiolabeling, an exponential phase culture of *E. coli* FB8-LysA Str^r was prepared in LB medium containing streptomycin (50 mg/ml) and kanamycin (25 mg/ml) and was used to inoculate (1:100) 500 ml of prewarmed M9 minimal media supplemented with threonine, methionine, and lysine (Sigma-Aldrich), each at 100 μg/ml. For radiolabeling, [^{3}H]*meso*-DAP (50 μCi/liter; Moraveck Inc.) was added. A nonlabeled control culture of *E. coli* FB8-LysA Str^r was prepared in parallel. Cultures were incubated overnight at 37°C with aeration. Final OD_{600} was ~2.0. Bacteria were washed five times by centrifugation at 4000*g*, 4°C and resuspended with cold PBS. Removal of free [^{3}H]*meso*-DAP was confirmed by scintillation counting of 1 ml of supernatant. To favor *E. coli* FB8-LysA Str^r intestinal colonization, streptomycin sulfate (5 g/liter) was added to the drinking water of mice, beginning 24 hours before gavage, and maintained until the end of the experiment. Mice were gavaged with ~10^{10} colony-forming units (CFU) of [^{3}H]*meso*-DAP–labeled *E. coli* FB8-LysA Str^r or the unlabeled control strain. Blood and tissues were collected 24 hours after gavage, as described below. To assess colonization efficacy, *E. coli* CFUs were quantified by serial dilution of fecal homogenates spotted onto LB agar containing streptomycin (50 mg/ml) and kanamycin (25 mg/ml) after overnight incubation at 37°C.

Processing and measurement of ^{14}C and ^{3}H from blood and brain homogenates

Profound anesthesia was induced in mice by intraperitoneal injection of ketamine (100 mg/kg; Imalgene1000, Boehringer-Ingelheim) and xylazine (8 mg/kg; Rompun 2%, Bayer). Blood collection was performed upon section of the inferior vena cava and was immediately followed by transcardial perfusion with 20 ml of PBS with a syringe connected to a 26G needle to eliminate blood from the circulatory system. The brain and small intestine were then extracted. Approximately 8 cm of small intestine proximal to the stomach ("duodenum" in figs. S3F and S5F) and 8 cm of small intestine proximal to the caecum ("ileum" in figs. S3F and S5F) were collected. Tissue weights were determined using a precision balance. Brain, duodenum, ileum, and blood (200 μl) were transferred into glass vials and dissolved with 2 ml of Solvable (PerkinElmer) overnight at 60°C, followed by color-bleaching by addition of 500 μl (for blood) or 200 μl (for tissues) of 30% hydrogen peroxide (Millipore). The tubes were incubated for 30 min at room temperature and then 30 min at 60°C. The solutions were then transferred to 20-ml HDPE scintillation vials with urea cap containing a polyethylene cone (Duran Wheaton Kimble) and combined with 10 ml of UltimaGold LTT scintillation cocktail (PerkinElmer). After a minimum period of 4 hours for temperature equilibration in the dark, the samples were analyzed using a Tri-Carb 3110 TR Liquid Scintillation Analyzer with QuantaSmart TriCarb LCS 3.00 software. ^{14}C was measured in the range 0.0 to 156 keV for 2 min and ^{3}H was measured in the range 2.0 to 18.6 keV for 5 min, both on the high-sensitivity setting.

Food consumption

For overnight food consumption, mice were single-housed around 7 p.m. and a pre-weighed amount of food pellets was added into the cage. The next day (or 2 days later for the 40-hour measurements) around 10 a.m., food pellets were weighed and the weight difference was noted as the food was consumed. Automated food consumption analysis was performed using an automated food dispenser (FED system) (*50*) that was adapted with minor modifications. Mice were habituated to the FED system a few days prior to the analysis. During the test period, mice were individually housed with the FED device for 3 days. Mouse food consumption was evaluated on day 2. A meal bout was defined as a continuous period of food intake (longest nonfeeding interval accepted during a meal bout <2 min). To evaluate food consumption after MDP gavage, mice were individually housed with a FED device in their home cage for 5 days. Day 0: FED in; day 1: MDP control gavage; day 4: MDP gavage; day 5: FED was removed, and mice were placed together in their original cage. Mice were gavaged (1 mg in 0.1 ml) at 5 p.m. and food consumption was analyzed during the four subsequent hours. Outliers were identified using Grubbs' method (α = 0.05) and then removed. Mice that ate less than 4 g in 40 hours were also excluded from the analysis. Whenever possible, behavioral measures were performed blinded to mouse genotype.

Nest building

Mice were individually housed around 7 p.m. The next day at around 11 a.m., mice received a new pre-weighed cotton nestlet. Four hours later, any cotton that was not unrolled was removed and weighed. Mice were placed together in their original cages after the test. Whenever possible, behavioral measures were performed blinded to mouse genotype.

Light-dark box test

A two-compartment box containing a dark chamber (black walls with upper lid) and a light chamber (plexiglass walls, no upper lid) was used. The chambers were connected by a 10-cm^2 door in the middle of the wall. Animals were placed in one corner of the light chamber facing the wall and were allowed to freely explore for 6 min. The time spent in the dark chamber and the number of transitions from one chamber to the other were annotated manually afterward by looking at recorded sessions.

Temperature analysis

Telemetric temperature probes (Anipill pills, BodyCap medical) were implanted subcutaneously. After general gas anesthesia (isoflurane 3% for induction, followed by 1.5% for maintenance; diluted in 98.5% oxygen at a flow rate of 1 liter/min) and analgesia (buprenorphine, Buprecare; 0.1 mg/kg, subcutaneous), the skin in the scapular region was shaved, sterilized, locally anesthetized (lidocaine, Xylovet, 2%), and incised for probe insertion in a subcutaneous chamber. After probe insertion and skin suturing, mice were given buprenorphine every 24 hours for the next 48 hours. Mice were allowed to recover for a week before any data were collected. Temperature was measured every 15 min.

Drugs used

β_3-Adrenergic agonist (CL 316,243; Sigma-Aldrich) was diluted in sterile water and injected intraperitoneally (500 ng per animal). MDP (L-D isomer, active; tlrl-mdp) and MDP control (MDPctr; L-L isomer, inactive; tlrl-mdpcl) were purchased from InvivoGen. MDP and MDPctr were diluted in LAL-water (10 mg/ml), and 1 mg of either MDP of MDPctr was administered by oral gavage with a sterile gavage needle. Mice were systematically habituated to oral gavage prior to the test.

Glucose and insulin measurements

Mice were fasted for 6 hours. Blood samples from the tail vein were collected using microvette tubes (Sarstedt 16444). For blood glucose measurements, samples were immediately analyzed with a glucose meter (Accu Check,

Roche). For insulin measurements, samples were immediately placed on ice and centrifuged at 14,000*g* for 10 min at 4°C as soon as possible. Serum was collected and transferred into a new clean tube. Insulin concentration was determined using an ELISA kit (Mercodia, Uppsala, Sweden) according to the manufacturer's instructions. Homeostatic Model Assessment for Insulin Resistance (HOMA-IR) was calculated by multiplying fasting plasma insulin (FPI) by fasting plasma glucose (FPG), then dividing by the constant 22.5; HOMA-IR = (FPI × FPG)/22.5.

iDISCO/ClearMap

Mice were single-housed around 7 p.m. and were gavaged the next day at 10 a.m. with MDP or MDPctr (1 mg in 0.1 ml of LAL water). Mice were perfused 3 hours after gavage. The iDISCO and ClearMap protocol was followed as detailed on the website http://idisco.info (*51*). Whole brains were stained with anti–c-Fos (rabbit; 1:1500; Synaptic Systems 226-003) and Alexa Fluor 647–conjugated donkey anti-rabbit IgG (1:1000; Thermo Fisher A-31573).

Stereotaxic surgery

Mice were deeply anesthetized after an intraperitoneal injection of a mixture of ketamine (Imalgene 1000, 150 mg/kg) and xylazine (Rompun 2%, 5 mg/kg) and injected subcutaneously with buprenorphine (Buprecare; 0.1 mg/kg). After their heads were shaved, the mice were placed in a stereotaxic apparatus. The skin was sterilized and local anesthesia (lidocaine, Xylovet, 2%) was injected subcutaneously before the skin was incised. A small craniotomy was performed and the viral solution was then injected into DMH and/or ARC through a glass micropipette attached to a nanoinjector system (Nanoject III; 100 nl per site; 1 nl/s) using stereotaxic coordinates (–1.7 mm anterior from bregma, 0.3 mm lateral, and at a depth of 5 mm for DMH and 5.9 mm for ARC from the skull surface at bregma). Recombinant AAV expressing GCaMP6f (AAV9.hSyn.Flex.GCaMP6f; ~10^{13} VG/ml), GCaMP7f (AAV9.hSyn.Flex.GCaMP7f; ~10^{13} VG/ml), Cre recombinase (AAV9.hSyn.Cre; ~5 × 10^{11} VG/ml), mCherry (AAV9.hSyn.Flex.mCherry; ~10^{13} VG/ml), and TdTomato (AAV9.CAG.Flex.Tdtomato; ~10^{13} VG/ml) were purchased from Addgene Vector Core.

Fiber photometry

A fiber photometry system was developed based on (*52*) with modifications. Immediately after viral injection in the right ARC nucleus, a single optical fiber (multimode, 430 μm in diameter, NA 0.50, LC zirconia ferrule, Doric Lenses) was implanted 100 μm above the virus injection site and fixed to the skull with a liquid bonding resin (Superbond, Sun Medical) and dental acrylic (Unifast). At least 3 weeks after injection, GCaMP was continuously excited using a 473-nm diode-pumped solid-state (DPSS) laser (output fiber intensity, 0.1 mW; Crystal Lasers) reflected on a dichroic filter (452 to 490 nm/505 to 800 nm, MD498 GFP, Thorlabs) and collimated into a 400-μm multimode optical fiber (NA, 0.50; Doric Lenses) with a convergent lens (f = 30 mm). The emitted fluorescence was collected in the same fiber and transmitted by the dichroic mirror, filtered (525 ± 19 nm) and focused on a NewFocus 2151-fW photoreceptor (Newport; DC mode). Reflected blue light along the light path was also measured with another amplified photodetector (PDA36A, Thorlabs) for monitoring light excitation and fiber coupling. Signals from both photodetectors were digitized at 5000 Hz by a digital-to-analog converter (DAC; Power 1401, CED), and recorded and analyzed using Spike2 software. For simultaneous dual-color photometry, a 473-nm DPSS laser and a 561-nm DPSS laser (output fiber intensity, 0.1 mW; Crystal Lasers) were reflected via a dichroic filter (MD498 GFP, Thorlabs) onto a 488/561-nm dual-edge dichroic beam splitter (Di01-R488/561, Semrock) and collimated into a 400-μm multimode optical fiber (NA, 0.50, Doric Lenses) with a convergent lens (f = 30 mm). The emitted fluorescence was collected in the same fiber and separated by a 510-nm dichroic mirror, filtered (525 ± 19 nm for GCaMP and 624 ± 75 nm for mCherry) and focused on two respective NewFocus 2151-fW photoreceptors (Newport; DC mode). Animals were habituated to the fiber optic patch cord (Doric Lenses) and to the connection procedure at least three times before the recording sessions. Recordings with low fluorescence signals (mean fluorescence <0.15 mV and/or spontaneous events in $\Delta F/F$ <1%) or displaying movement/laser artifacts were discarded from the analysis. Animals were brought into the recording room and singly housed around 7 p.m. the night before the recording session. For the analysis under fasting conditions, food was removed when mice were brought into the recording room. For the analysis of steady-state (fed) conditions, mice were fed ad libitum at all times except during the recording session. Water was available in both conditions. On the following day at around 9 a.m., food was removed for the fed state and mice were recorded in the home cage for ~60 min to establish baseline signal. Then mice were either presented with a food pellet and recorded for ~60 min or gavaged with MDP or MDPctr (1 mg in 0.1 ml) and recorded for ~150 min. After this period, gavaged mice were presented with a food pellet and recorded for an additional 30-min period. For signal analysis of spontaneous activity, raw fluorescence signals were normalized ($\Delta F/F$) to the mean fluorescence (20-s sliding window) and then smoothened (0.05-s window). The standard deviation of the $\Delta F/F$ baseline signal was calculated and all the calcium-positive transients above 1.5 SD were automatically extracted with custom scripts (Spike2, CED, UK; minimum time interval between events of 5 s). The frequency of the sorted spontaneous events was then reported as the number of events per minute and normalized to the baseline activity before gavage. Nonspecific modulation of fluorescence signals was examined using simultaneous dual-color photometry of the calcium-sensitive fluorescent reporter GCaMP7f and the calcium-insensitive fluorescent report mCherry. mCherry traces were qualitatively different from GCaMP, with no noticeable fluctuation with the same laser power configuration (peak mCherry $\Delta F/F$ < 1%) and no concomitant/correlated events on both channels. To quantify the absence of synchronous artifacts in both GCaMP and mCherry channels, we sorted GCaMP events and calculated a GCaMP event-triggered mCherry trace average. We also calculated for each sorted GCaMP event the mean $\Delta F/F$ (1-s window around event) and the peak amplitude in both GCaMP and mCherry channels. With these quantifications, we did not observe any correlation between GCaMP and mCherry channels, ruling out any significant contribution of nonspecific events in GCaMP signals. Animals in which post hoc histological examination showed that viral injection or implanted optic fiber were not in the correct location were excluded from analysis. Investigators were not blinded to the group identity, but automatic analysis routines were performed with the same script executed for all experimental groups.

Patch-clamp recordings

Vgat^{cre}*Nod2*^{GFP} mice previously injected with a Cre-dependent Td-tomato virus in the ARC were deeply anesthetized with intraperitoneal injection of ketamine (100 mg/kg) and xylazine (10 mg/kg) and swiftly decapitated. The brain was rapidly dissected and placed in ice-cold artificial cerebrospinal fluid (ACSF) containing 124 mM NaCl, 3 mM KCl, 2.8 mM $MgSO_4$, 26 mM $NaHCO_3$, 1.25 mM $NaHPO_4$, 20 mM glucose, and 0.5 mM CaCl [~310 mOsm, pH 7.3 when bubbled with a mixture of 95% O_2 and 5% (vol/vol) CO_2] (all chemicals from Sigma-Aldrich). Coronal slices (250 μm thick) of the whole brain were placed in bubbling ACSF in a warm bath at 35°C for 30 min and then allowed to recover at room temperature (i.e., 22° ± 1°C). For whole-cell recordings, individual slices were placed in a chamber mounted on a Zeiss Axioskop upright microscope and continuously perfused (1.5 ml/min) with 30°C ACSF (Warner Instrument inline heater) containing 124 mM NaCl, 3 mM KCl, 1.3 mM $MgSO_4$, 26 mM $NaHCO_3$, 1.25 mM $NaHPO_4$, 20 mM glucose, and 2 mM CaCl.

Slices were visualized using a 40× water immersion objective (Zeiss Achroplan 440090; NA 0.75; no aberration correction; working distance 2.1 mm). We performed whole-cell patch-clamp recordings from visually targeted cells expressing both Td-Tomato and GFP in the ARC. Patch pipettes were pulled from borosilicate glass (OD 1.5 mm, ID 0.86 mm; P-97 Flaming/Brown micropipette puller, Sutter Instruments). These pipettes were filled with a solution containing a mix of MDP or MDPctrl (20.3 μM), biocytin (0.15%, Sigma-Aldrich), and potassium methanesulfonate–based solution (126 mM K-methansesulfonate, 3.6 mM KCl, 3.6 mM NaCl, 9 mM Na-Hepes, 10 mM Cs-EGTA, 0.3 mM GTP, and 4 mM Mg-ATP, 280 to 290 mOsm, pH 7.3). The solution passively diffused from the patch pipette into the recorded cell after membrane rupture in the whole-cell configuration. Recordings were obtained via an Axon Multiclamp 700B. Data were acquired using Elphy software (G. Sadoc, CNRS, Gif-sur-Yvette, France). Recordings were performed in the current-clamp mode. Constant current injection was used to hold membrane potential near −70 mV. Hyperpolarizing current steps (20 pA) were used to monitor input resistance and incremental depolarizing current steps (500 ms, 5 pA, up to 120 pA) were used to elicit action potentials. The number of action potentials elicited was analyzed for each depolarizing step. This protocol was applied just after membrane rupture and was repeated after 10, 20, and 30 min of recording. Only cells showing sustained trains of action potentials at the initial time point were considered for further analysis. Membrane potentials are corrected for a measured liquid junction potential of 10 mV. The rheobase value was the minimal depolarizing current pulse sufficient to trigger an action potential. Recordings were analyzed with Elphy software. After recording, brain slices were fixed overnight in 10% PFA and then washed three times in 1× PBS. Slices were stained with anti-GFP as described above and incubated with streptavidin (1:1000) in 1× PBS with 0.3% Triton X-100 for 30 min at room temperature. Slices were then washed and mounted. Images were captured with 40× Zeiss lens (LD Plan-Neofluar 40×/0.6 Korr Ph 2M27) on a confocal laser-scanning microscope (LSM 700, Zeiss).

Long-term antibiotic treatment

For mice injected with a Cre-expressing virus, the antibiotic treatment was initiated 1 week before the stereotaxic injections and lasted for 13 weeks. Mice received a cocktail of antibiotics in their drinking water, which contained streptomycin (2.5 mg/ml), ampicillin (1 mg/ml),vancomycin (0.5 mg/ml), and metronidazole (0.5 mg/ml; all from Sigma-Aldrich). The antibiotic-containing drinking water was changed once a week and the bottles were covered to avoid light exposure. Mouse weights were measured and feces were collected weekly.

Fecal DNA extraction and bacterial qPCR

Fecal DNA was extracted using the FastDNA spin kit (MP biomedical), following the recommended protocol. Five nanograms of DNA were used to perform the qPCR using SybrGreen (BioRad). qPCR conditions and primers were based on a previous publication (*53*). The following sets of primers were used: pan-bacteria: 5′-GCAGGCCTAACACATGCAAGTC-3′ (forward) and 5′-CTGCTGCCTCCCGTAGGAGT-3′ (reverse); Actinobacteria: 5′-TACGGCCGCAAGGCT-3′ (forward) and 5′-TCRTCCCCACCTTCCTCCG-3′ (reverse); Gammaproteobacteria: 5′-TCGTCAGCTCGTGTYGTGA-3′ (forward) and 5′-CGTAAGGGCCATGATG-3′ (reverse); Bacteroides: 5′-CRAACAGGATTAGATACCCT-3′ (forward) and 5′-GGTAAGGTTCCTCGCGTAT-3′ (reverse); and Firmicutes: 5′-TGAAACTYAAAGGAATTGACG-3′ (forward) and 5′-ACCATGCACCTGTC-3′ (reverse).

The pan-bacteria Ct value for each mouse was normalized to the mean Ct obtained from feces collected just before ABX treatment was begun. Bacterial group Ct values were normalized to the Pan-bacteria Ct of the respective mouse and day.

Statistics

Statistical analysis was performed using commercial analysis software (GraphPad Prism) and with R 3.6.3 (www.R-project.org), with $P \leq 0.05$ considered significant. Each experiment was replicated at least twice. For the "nest-test" analysis (Figs. 2G and 5L and fig. S4E), due to the distribution of values associated with the percentage of unrolled cotton, we chose to compare discretized versions of the corresponding distributions. Values were binned as: class 1, 0 to 25%; class 2, 25 to 50%; class 3, 50 to 75%; class 4, 75 to 100%. The distribution of events into these different classes was then compared among the groups using a Fisher's exact test. The graphs show the data for each mouse and not the class distribution for a better appreciation of the raw values. For Fig. 5N analysis, the time is the driving force on the two-way analysis of variance (ANOVA); when looking into the contrast, one can appreciate that the top driver of this force is the control group ($P = 0.0096$), whereas the floxed group is not significant ($P = 0.085$). For Fig. 1F and fig. S3, B to E, the control group was not considered for the statistical analysis because this group did not receive any radiolabeled PG-derivate. In supplementary figures, males and females are shown in the same graph only for a better graphical representation. However, for statistical analysis, they were analyzed as two independent variables.

REFERENCES AND NOTES

1. L. H. Morais, H. L. Schreiber4th, S. K. Mazmanian, The gut microbiota-brain axis in behaviour and brain disorders. *Nat. Rev. Microbiol.* **19**, 241–255 (2021). doi: 10.1038/s41579-020-00460-0; pmid: 33093662
2. J. K. Nicholson *et al*., Host-gut microbiota metabolic interactions. *Science* **336**, 1262–1267 (2012). doi: 10.1126/science.1223813; pmid: 22674330
3. J. F. Cryan, T. G. Dinan, Mind-altering microorganisms: The impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* **13**, 701–712 (2012). doi: 10.1038/nrn3346; pmid: 22968153
4. E. Y. Hsiao *et al*., Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* **155**, 1451–1463 (2013). doi: 10.1016/j.cell.2013.11.024; pmid: 24315484
5. J. A. Bravo *et al*., Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16050–16055 (2011). doi: 10.1073/pnas.1102999108; pmid: 21876150
6. S. A. Buffington *et al*., Dissecting the contribution of host genetics and the microbiome in complex behaviors. *Cell* **184**, 1740–1756.e16 (2021). doi: 10.1016/j.cell.2021.02.009; pmid: 33705688
7. C. A. Thaiss *et al*., Microbiota Diurnal Rhythmicity Programs Host Transcriptome Oscillations. *Cell* **167**, 1495–1510.e12 (2016). doi: 10.1016/j.cell.2016.11.003; pmid: 27912059
8. S. E. Girardin *et al*., Peptidoglycan molecular requirements allowing detection by Nod1 and Nod2. *J. Biol. Chem.* **278**, 41702–41708 (2003). doi: 10.1074/jbc.M307198200; pmid: 12871942
9. T. Arentsen *et al*., The bacterial peptidoglycan-sensing molecule Pglyrp2 modulates brain development and behavior. *Mol. Psychiatry* **22**, 257–266 (2017). doi: 10.1038/mp.2016.182; pmid: 27843150
10. A. Masuzzo *et al*., Peptidoglycan-dependent NF-κB activation in a small subset of brain octopaminergic neurons controls female oviposition. *eLife* **8**, e50559 (2019). doi: 10.7554/eLife.50559; pmid: 31661076
11. J. Royet, D. Gupta, R. Dziarski, Peptidoglycan recognition proteins: Modulators of the microbiome and inflammation. *Nat. Rev. Immunol.* **11**, 837–851 (2011). doi: 10.1038/nri3089; pmid: 22076558
12. S. E. Girardin *et al*., Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J. Biol. Chem.* **278**, 8869–8872 (2003). doi: 10.1074/jbc.C200651200; pmid: 12527755
13. L. Cheng *et al*., NOD2 promotes dopaminergic degeneration regulated by NADPH oxidase 2 in 6-hydroxydopamine model of Parkinson's disease. *J. Neuroinflammation* **15**, 243 (2018). doi: 10.1186/s12974-018-1289-z; pmid: 30157869
14. A. Fani Maleki *et al*., Muramyl dipeptide-mediated immunomodulation on monocyte subsets exerts therapeutic effects in a mouse model of Alzheimer's disease. *J. Neuroinflammation* **17**, 218 (2020). doi: 10.1186/s12974-020-01893-3; pmid: 32698829
15. J. Oliveira *et al*., Genetic association between a 'standing' variant of NOD2 and bipolar disorder. *Immunobiology* **219**, 766–771 (2014). doi: 10.1016/j.imbio.2014.06.003; pmid: 25053139
16. J. E. van Schijndel *et al*., Three-cohort targeted gene screening reveals a non-synonymous TRKA polymorphism associated with schizophrenia. *J. Psychiatr. Res.* **43**, 1195–1199 (2009). doi: 10.1016/j.jpsychires.2009.04.006; pmid: 19435634
17. Q. Ma *et al*., P268S in NOD2 associates with susceptibility to Parkinson's disease in Chinese population. *Behav. Brain Funct.* **9**, 19 (2013). doi: 10.1186/1744-9081-9-19; pmid: 23651603
18. M. J. Pabst, S. Beranova-Giorgianni, J. M. Krueger, Effects of muramyl peptides on macrophages, monokines, and sleep. *Neuroimmunomodulation* **6**, 261–283 (1999). doi: 10.1159/000026384; pmid: 10393513
19. E. Denou *et al*., Defective NOD2 peptidoglycan sensing promotes diet-induced inflammation, dysbiosis, and insulin resistance. *EMBO Mol. Med.* **7**, 259–274 (2015). doi: 10.15252/emmm.201404169; pmid: 25666722
20. J. F. Cavallari *et al*., Muramyl Dipeptide-Based Postbiotics Mitigate Obesity-Induced Insulin Resistance via IRF4. *Cell Metab.* **25**, 1063–1074.e3 (2017). doi: 10.1016/j.cmet.2017.03.021; pmid: 28434881
21. W. Jiang *et al*., Recognition of gut microbiota by NOD2 is essential for the homeostasis of intestinal intraepithelial lymphocytes. *J. Exp. Med.* **210**, 2465–2476 (2013). doi: 10.1084/jem.20122490; pmid: 24062413

22. R. D. Lisk, R. A. Pretlow 3rd, S. M. Friedman, Hormonal stimulation necessary for elicitation of maternal nest-building in the mouse (Mus musculus). *Anim. Behav.* **17**, 730–737 (1969). doi: 10.1016/S0003-3472(69)80020-5; pmid: 5378004
23. S. Suyama, T. Yada, New insight into GABAergic neurons in the hypothalamic feeding regulation. *J. Physiol. Sci.* **68**, 717–722 (2018). doi: 10.1007/s12576-018-0622-8; pmid: 30003408
24. S. F. Morrison, Central neural control of thermoregulation and brown adipose tissue. *Auton. Neurosci.* **196**, 14–24 (2016). doi: 10.1016/j.autneu.2016.02.010; pmid: 26924538
25. Y. Chen, Y.-C. Lin, T.-W. Kuo, Z. A. Knight, Sensory detection of food rapidly modulates arcuate feeding circuits. *Cell* **160**, 829–841 (2015). doi: 10.1016/j.cell.2015.01.033; pmid: 25703096
26. Z. Huang *et al*., Antibody neutralization of microbiota-derived circulating peptidoglycan dampens inflammation and ameliorates autoimmunity. *Nat. Microbiol.* **4**, 766–773 (2019). doi: 10.1038/s41564-019-0381-1; pmid: 30833732
27. F. De Vadder *et al*., Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell* **156**, 84–96 (2014). doi: 10.1016/j.cell.2013.12.016; pmid: 24412651
28. J. Breton *et al*., Gut commensal E. coli proteins activate host satiety pathways following nutrient-induced bacterial growth. *Cell Metab.* **23**, 324–334 (2016). doi: 10.1016/j.cmet.2015.10.017; pmid: 26621107
29. T. Arentsen, R. Khalid, Y. Qian, R. Diaz Heijtz, Sex-dependent alterations in motor and anxiety-like behavior of aged bacterial peptidoglycan sensing molecule 2 knockout mice. *Brain Behav. Immun.* **67**, 345–354 (2018). doi: 10.1016/j.bbi.2017.09.014; pmid: 28951252
30. M. M. Pusceddu *et al*., Nod-like receptors are critical for gut-brain axis signalling in mice. *J. Physiol.* **597**, 5777–5797 (2019). doi: 10.1113/JP278640; pmid: 31652348
31. G. Pei *et al*., Cellular stress promotes NOD1/2-dependent inflammation via the endogenous metabolite sphingosine-1-phosphate. *EMBO J.* **40**, e106272 (2021). doi: 10.15252/embj.2020106272; pmid: 33942347
32. L. M. Cox, H. Abou-El-Hassan, A. H. Maghzi, J. Vincentini, H. L. Weiner, The sex-specific interaction of the microbiome in neurodegenerative diseases. *Brain Res.* **1724**, 146385 (2019). doi: 10.1016/j.brainres.2019.146385; pmid: 31419428
33. M.-C. Audet, Stress-induced disturbances along the gut microbiota-immune-brain axis and implications for mental health: Does sex matter? *Front. Neuroendocrinol.* **54**, 100772 (2019). doi: 10.1016/j.yfrne.2019.100772; pmid: 31302116
34. K. M. Culbert, C. L. Sisk, K. L. Klump, A Narrative Review of Sex Differences in Eating Disorders: Is There a Biological Basis? *Clin. Ther.* **43**, 95–111 (2021). doi: 10.1016/j.clinthera.2020.12.003; pmid: 33375999
35. A. Franceschini, L. Fattore, Gender-specific approach in psychiatric diseases: Because sex matters. *Eur. J. Pharmacol.* **896**, 173895 (2021). doi: 10.1016/j.ejphar.2021.173895; pmid: 33508283
36. M. S. Thion *et al*., Microbiome Influences Prenatal and Adult Microglia in a Sex-Specific Manner. *Cell* **172**, 500–516.e16 (2018). doi: 10.1016/j.cell.2017.11.042; pmid: 29275859
37. F. Kühn *et al*., Intestinal alkaline phosphatase targets the gut barrier to prevent aging. *JCI Insight* **5**, e134049 (2020). doi: 10.1172/jci.insight.134049; pmid: 32213701
38. E. Jašarević, K. E. Morrison, T. L. Bale, Sex differences in the gut microbiome-brain axis across the lifespan. *Philos. Trans. R. Soc. London Ser. B* **371**, 20150122 (2016). doi: 10.1098/rstb.2015.0122; pmid: 26833840
39. Z. S. Ma, W. Li, How and Why Men and Women Differ in Their Microbiomes: Medical Ecology and Network Analyses of the Microgenderome. *Adv. Sci.* **6**, 1902054 (2019). doi: 10.1002/advs.201902054; pmid: 31832327
40. S. L. Padilla *et al*., AgRP to Kiss1 neuron signaling links nutritional state and fertility. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2413–2418 (2017). doi: 10.1073/pnas.1621065114; pmid: 28196880
41. G. Neal-Perry, E. Nejat, C. Dicken, The neuroendocrine physiology of female reproductive aging: An update. *Maturitas* **67**, 34–38 (2010). doi: 10.1016/j.maturitas.2010.04.016; pmid: 20570066
42. F. Mauvais-Jarvis, D. J. Clegg, A. L. Hevener, The role of estrogens in control of energy balance and glucose homeostasis. *Endocr. Rev.* **34**, 309–338 (2013). doi: 10.1210/er.2012-1055; pmid: 23460719
43. M. E. Griffin, C. W. Hespen, Y. C. Wang, H. C. Hang, Translation of peptidoglycan metabolites into immunotherapeutics. *Clin. Transl. Immunol.* **8**, e1095 (2019). doi: 10.1002/cti2.1095; pmid: 31798878
44. D. Kim *et al*., Nod2-mediated recognition of the microbiota is critical for mucosal adjuvant activity of cholera toxin. *Nat. Med.* **22**, 524–530 (2016). doi: 10.1038/nm.4075; pmid: 27064448
45. F. Barreau *et al*., CARD15/NOD2 is required for Peyer's patches homeostasis in mice. *PLOS ONE* **2**, e523 (2007). doi: 10.1371/journal.pone.0000523; pmid: 17565376
46. E. Casanova *et al*., A CamKIIα iCre BAC allows brain-specific gene inactivation. *Genesis* **31**, 37–42 (2001). doi: 10.1002/gene.1078; pmid: 11668676
47. R. Wheeler, F. Veyrier, C. Werts, I. G. Boneca, in *Glycoscience: Biology and Medicine*, N. Taniguchi *et al.*, Eds. (Springer, 2015), pp. 737–747.
48. D. Mengin-Lecreulx, E. Siegel, J. van Heijenoort, Variations in UDP-N-acetylglucosamine and UDP-N-acetylmuramyl-pentapeptide pools in Escherichia coli after inhibition of protein synthesis. *J. Bacteriol.* **171**, 3282–3287 (1989). doi: 10.1128/jb.171.6.3282-3287.1989; pmid: 2656647
49. F. B. Wientjes, C. L. Woldringh, N. Nanninga, Amount of peptidoglycan in cell walls of gram-negative bacteria. *J. Bacteriol.* **173**, 7684–7691 (1991). doi: 10.1128/jb.173.23.7684-7691.1991; pmid: 1938964
50. K. P. Nguyen, T. J. O'Neal, O. A. Bolonduro, E. White, A. V. Kravitz, Feeding Experimentation Device (FED): A flexible open-source device for measuring feeding behavior. *J. Neurosci. Methods* **267**, 108–114 (2016). doi: 10.1016/j.jneumeth.2016.04.003; pmid: 27060385
51. N. Renier *et al*., Mapping of Brain Activity by Automated Volume Analysis of Immediate Early Genes. *Cell* **165**, 1789–1802 (2016). doi: 10.1016/j.cell.2016.05.007; pmid: 27238021
52. L. A. Gunaydin *et al*., Natural neural projection dynamics underlying social behavior. *Cell* **157**, 1535–1551 (2014). doi: 10.1016/j.cell.2014.05.017; pmid: 24949967
53. T. Bacchetti De Gregoris, N. Aldred, A. S. Clare, J. G. Burgess, Improvement of phylum- and class-specific primers for real-time PCR quantification of bacterial taxa. *J. Microbiol. Methods* **86**, 351–356 (2011). doi: 10.1016/j.mimet.2011.06.010; pmid: 21704084

ACKNOWLEDGMENTS

We thank S. Saha, E. de Launoit, and T. Topilko for their help in conducting and/or analyzing experiments; G. Chevalier for his preliminary work on *Nod2*^{GFP} mice; and D. Mucida and G. M. Silva for their insightful discussions and critical reading and editing of the manuscript. We would also like to thank P. Campagne from the Hub of Bioinformatics and Biostatistics of the Institut Pasteur for his advice on the statistical tests used in this work and the members of the Animalerie Centrale of the Institut Pasteur for maintenance and care of the mice used in this work. We also thank the members of the Microenvironment & Immunity Unit, the laboratory for Perception & Memory, the Stroma, Inflammation & Tissue Repair Unit, and the Biology and Genetics of the Bacterial Cell Wall Unit, for their support and feedback. **Funding:** Institutional support was provided by Institut Pasteur, CNRS, and Inserm. Also supported by Pasteur-Roux postdoctoral fellowships from the Institut Pasteur (I.G.), a Human Frontier Science Program fellowship (I.G.), Agence Nationale de la Recherche grant ANR-16-CE15-0021 (I.G.B., P.-M.L., G.C. and G.E.), DFG CRC1182 project C2 (P.R.), and life insurance company "AG2R-La-Mondiale" (P.-M.L.). **Author contributions:** Conceptualization: I.G. Formal analysis: I.G., A.V.-P. Funding acquisition: I.G., G.L., I.G.B., G.E., P.-M.L. Investigation: I.G., G.L., N.R., A.V.-P., R.W., C.M., A.N., S.D., B.P., F.V. Methodology: I.G., G.L., N.R., S.W., A.V.-P., A.N., R.W. Project administration: I.G. Resources: S.H., P.R., I.G.B. Supervision: G.L., I.G.B., G.E., P.-M.L. Validation: I.G., N.R., A.V.-P., G.L., R.W. Visualization: I.G., A.V.-P. Writing–original draft: I.G. Writing–review and editing: I.G., G.L., N.R., I.G.B., G.E., P.-M.L., R.W., P.R. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj3986
Figs. S1 to S8
Tables S1 and S2
MDAR Reproducibility Checklist

11 May 2021; accepted 25 February 2022
10.1126/science.abj3986

RESEARCH ARTICLE SUMMARY

SIGNAL TRANSDUCTION

Inhibition of nonalcoholic fatty liver disease in mice by selective inhibition of mTORC1

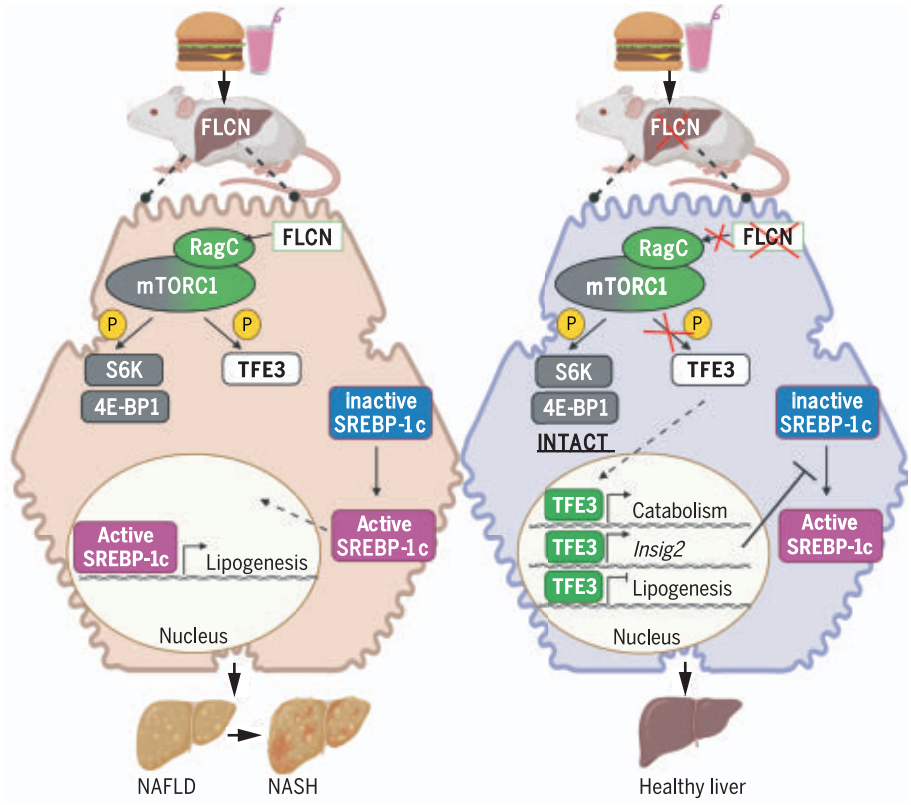
Bridget S. Gosis, Shogo Wada, Chelsea Thorsheim, Kristina Li, Sunhee Jung, Joshua H. Rhoades, Yifan Yang, Jeffrey Brandimarto, Li Li, Kahealani Uehara, Cholsoon Jang, Matthew Lanza, Nathan B. Sanford, Marc R. Bornstein, Sunhye Jeong, Paul M. Titchenell, Sudha B. Biddinger, Zoltan Arany*

INTRODUCTION: As many as 100 million people in the US have nonalcoholic fatty liver disease (NAFLD), characterized by increased liver lipid accumulation, which often leads to hepatocyte injury and fibrosis, characteristics of nonalcoholic steatohepatitis (NASH). NASH in turn can progress to cirrhosis and hepatocellular carcinoma. There are currently no US Food and Drug Administration–approved therapies for NAFLD or NASH. NAFLD occurs when there is disequilibrium between the processes of hepatic lipid synthesis and consumption. The nutrient sensor mechanistic target of rapamycin complex 1 (mTORC1) regulates several of these pathways. mTORC1 is thus an attractive target to modulate lipid homeostasis in the liver. However, mTORC1 also regulates numerous other cellular pathways, and blunting of mTORC1 modulation can lead to unexpected feedback loops and unwanted effects.

RATIONALE: We hypothesized that selective modulation of hepatic mTORC1 signaling could benefit liver lipid metabolism and prevent NAFLD. In non-liver cell types, the protein folliculin (FLCN) has been shown to confer substrate specificity to mTORC1. Deletion of FLCN inhibits mTORC1-mediated phosphorylation of the transcription factor E3/B (TFE3/B) family of transcription factors, without affecting mTORC1-driven phosphorylation of its canonical substrates ribosomal protein S6 kinase beta-1 (S6K1) and eukaryotic translation initiation factor 4E–binding protein 1 (4E-BP1). Unphosphorylated TFE3 translocates to the nucleus and activates genes that promote lysosomal biogenesis, mitochondrial biogenesis, and oxidative metabolism. We reasoned that suppression of FLCN in the liver might promote fatty acid oxidation and lipid clearance without untoward effects of generalized mTORC1 inhibition.

RESULTS: Hepatocyte-specific genetic deletion of *Flcn* in adult mice selectively inhibited mTORC1-mediated cytoplasmic sequestration of TFE3, with little effect on other mTORC1 targets, including S6K, 4E-BP1, and Lipin1. Hepatocyte loss of *Flcn* protected mice from both NAFLD and NASH and partially reversed these processes when already established. The protection against NAFLD and NASH required TFE3, which activated lipid clearance. Unleashed TFE3 additionally suppressed de novo lipogenesis. The latter was mediated in part by TFE3-mediated induction of insulin-induced gene 2 (*Insig2*) to inhibit proteolytic activation of sterol regulatory element–binding protein-1c (SREBP-1c), a critical lipogenic transcription factor.

CONCLUSION: Our data establish FLCN as a critical regulator of lipid homeostasis in the liver. *Flcn* deletion affords selective inhibition of mTORC1, leading to nuclear translocation and activation of the transcription factor TFE3, which coordinates hepatic lipid metabolic pathways to protect against NAFLD and NASH in mice. Thus, our data reveal FLCN as a promising target for the treatment of NAFLD and NASH. The data also illuminate previously published and seemingly conflicting data, which likely reflected different effects on each arm of mTORC1 signaling. There have been numerous attempts by many to develop disease-specific treatments for NAFLD and NASH, thus far without success. A recurrent problem has been the many compensatory responses by the liver to targeting any one pathway; for example, inhibitors of acetyl–coenzyme A carboxylase led to compensatory activation of SREBP-1c and consequent hyperlipidemia. Targeting FLCN is thus particularly attractive, in that loss of FLCN simultaneously and favorably affects multiple aspects of hepatic lipid homeostasis, including promoting fatty acid oxidation and lysosomal biogenesis and inhibiting de novo lipogenesis. ■



Deletion of *Flcn* in the liver protects mice from NAFLD and NASH through selective suppression of mTORC1. Diets high in fat, carbohydrates, and cholesterol lead to NAFLD and NASH. When *Flcn* is simultaneously deleted, mTORC1 is selectively inhibited, preserving phosphorylation of canonical substrates S6K and 4E-BP1 while blocking phosphorylation of the transcription factor TFE3. Unphosphorylated TFE3 is released to the nucleus, where it activates lipid catabolism genes while suppressing de novo lipogenesis genes. [Image created using Biorender]

The list of author affiliations is available in the full article online.
*Corresponding author. Email: zarany@pennmedicine.upenn.edu
Cite this article as B. S. Gosis *et al.*, *Science* **376**, eabf8271 (2022). DOI: 10.1126/science.abf8271

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf8271

RESEARCH ARTICLE

SIGNAL TRANSDUCTION

Inhibition of nonalcoholic fatty liver disease in mice by selective inhibition of mTORC1

Bridget S. Gosis[1], Shogo Wada[1], Chelsea Thorsheim[1], Kristina Li[1], Sunhee Jung[2], Joshua H. Rhoades[1,3,4], Yifan Yang[1], Jeffrey Brandimarto[1], Li Li[1], Kahealani Uehara[5,6], Cholsoon Jang[2], Matthew Lanza[7], Nathan B. Sanford[1], Marc R. Bornstein[1], Sunhye Jeong[1], Paul M. Titchenell[5,6], Sudha B. Biddinger[8], Zoltan Arany[1]*

Nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH) remain without effective therapies. The mechanistic target of rapamycin complex 1 (mTORC1) pathway is a potential therapeutic target, but conflicting interpretations have been proposed for how mTORC1 controls lipid homeostasis. We show that selective inhibition of mTORC1 signaling in mice, through deletion of the RagC/D guanosine triphosphatase–activating protein folliculin (FLCN), promotes activation of transcription factor E3 (TFE3) in the liver without affecting other mTORC1 targets and protects against NAFLD and NASH. Disease protection is mediated by TFE3, which both induces lipid consumption and suppresses anabolic lipogenesis. TFE3 inhibits lipogenesis by suppressing proteolytic processing and activation of sterol regulatory element–binding protein-1c (SREBP-1c) and by interacting with SREBP-1c on chromatin. Our data reconcile previously conflicting studies and identify selective inhibition of mTORC1 as a potential approach to treat NASH and NAFLD.

As many as 100 million people in the US have nonalcoholic fatty liver disease (NAFLD), characterized by increased liver lipid accumulation (*1*), which often leads to hepatocyte injury and fibrosis, characteristics of nonalcoholic steatohepatitis (NASH) (*2*, *3*). NASH can, in turn, progress to cirrhosis and hepatocellular carcinoma (*1*, *4*). To date, there is no US Food and Drug Administration–approved therapy for NAFLD or NASH (*4*). NAFLD stems from a disequilibrium between hepatic lipid flux processes (*5*). Attempts at therapeutic modulation of any one of these processes often lead to feedback modulation of the others, limiting efficacy or causing unwanted side effects. Identifying therapeutic nodes that strike the proper balance has therefore been challenging.

The mechanistic target of rapamycin (mTOR) pathway is a critical nutrient-sensing pathway in many cell types, including hepatocytes (*6*). The mechanistic target of rapamycin complex 1 (mTORC1), nucleated around the adaptor protein Raptor, activates anabolic pathways, such as lipid and protein synthesis, and inhibits catabolic pathways, such as autophagy and oxidative metabolism (*6*, *7*). For this reason, mTORC1 has been studied as an attractive target to modulate lipid homeostasis in the liver, but its role in this process remains unclear, and studies point to seemingly opposing conclusions. Deletion of *Raptor* in hepatocytes suppresses de novo lipogenesis (DNL) through inhibition of the lipogenic transcription factor sterol regulatory element–binding protein-1c (SREBP-1c) and protected mice from liver steatosis (*8*). These results were consistent with an anabolic role of mTORC1 activity (*8*) and with the requirement of mTORC1 for activation of SREBP-1c by the protein kinase Akt (*9*). However, different studies reported that hepatocyte deletion of the TSC complex subunit 1 gene (*Tsc1*) (*10*, *11*), and therefore activation of mTORC1, also protected mice from liver steatosis. Other groups have observed increased steatosis (*12*) and liver injury (*13*) in mice lacking *Raptor* in the liver. Thus, the role of mTORC1 in liver steatosis remains poorly understood.

Substrate specificity can be conferred to mTORC1 by the protein folliculin (FLCN) (*14*–*16*). Heterozygous loss-of-function mutations in the *FLCN* gene cause Birt-Hogg-Dubé syndrome, characterized by benign skin fibrofolliculomas, lung cysts, and, in a subset of patients, renal cell carcinoma caused by loss of heterozygosity (*17*). FLCN is a guanosine triphosphatase (GTPase)–activating protein (GAP) for the GTPases RagC and RagD, which, in their guanosine diphosphate–bound state, promote activation of mTORC1 (*18*). Deletion of *Flcn* in various cell types inhibits mTORC1-mediated phosphorylation of one set of targets, the transcription factor E3/B (TFE3/B) family of transcription factors, without affecting mTORC1-driven phosphorylation of its canonical substrates ribosomal protein S6 kinase beta-1 (S6K1) and eukaryotic translation initiation factor 4E–binding protein 1 (4E-BP1) (*14*–*16*). As a result, modulation of FLCN allows separation of these two pathways downstream of mTORC1 (Fig. 1, A and B). Unphosphorylated TFE3 translocates to the nucleus and activates genes that promote lysosomal biogenesis, mitochondrial biogenesis, and oxidative metabolism (*15*, *19*–*22*). Thus, in adipose tissue, for example, FLCN deletion promotes beiging and thermogenesis (*14*, *23*), without suppressing S6K-mediated anabolism or causing lipodystrophy (*14*).

We reasoned that suppression of FLCN in the liver might promote fatty acid oxidation and lipid clearance without untoward effects of generalized mTORC1 inhibition. We found that deletion of *Flcn* in hepatocytes promoted oxidative metabolism and lysosomal biogenesis. Unexpectedly, we also found that deletion of *Flcn* in hepatocytes suppressed SREBP-1c activity and de novo lipogenesis through multiple mechanisms. In addition, activation of canonical mTORC1 signaling triggered a feedback loop that suppressed noncanonical FLCN-mTORC1 activity on TFE3, thereby in part explaining discordant findings reported in the literature. Ultimately, suppression of liver FLCN led to protection against, and reversal of, NAFLD and NASH.

Liver FLCN selectively promotes mTORC1-mediated cytoplasmic sequestration of TFE3, without affecting canonical mTORC1 signaling

Mice with liver-specific deletion of *Flcn* (LiFKO mice) were generated by infecting $Flcn^{lox/lox}$ mice with hepatotropic adeno-associated virus serotype 8 (AAV8) expressing Cre recombinase driven by the hepatocyte-specific thyroxine-binding globulin (TBG) promoter (Fig. 1C). FLCN content in LiFKO livers was decreased by more than 90% (Fig. 1D). TFE3 protein was enriched in the nuclei of cells from livers lacking *Flcn* (Fig. 1E), as was a slower migrating form of TFE3, as previously observed with *Flcn* deletion in different cell types (*15*, *22*). Similar experiments in mice lacking *Raptor* in hepatocytes (*12*) demonstrated comparable accumulation of nuclear TFE3 (Fig. 1F and fig. S1A). Thus, hepatic FLCN appears to activate mTORC1 and promote retention of TFE3 in the cytoplasm.

Phosphorylation of canonical mTORC1 targets, including S6 and 4E-BP1, was mostly unaffected in livers of LiFKO mice, whether fed normal chow (Fig. 1G) or a NAFLD-inducing

[1]Cardiovascular Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [2]Department of Biological Chemistry, University of California, Irvine, CA, USA. [3]Institute for Biomedical Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [4]School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, USA. [5]Institute for Diabetes, Obesity, and Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [6]Department of Physiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [7]Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, USA. [8]Division of Endocrinology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA.
*Corresponding author. Email: zarany@pennmedicine.upenn.edu

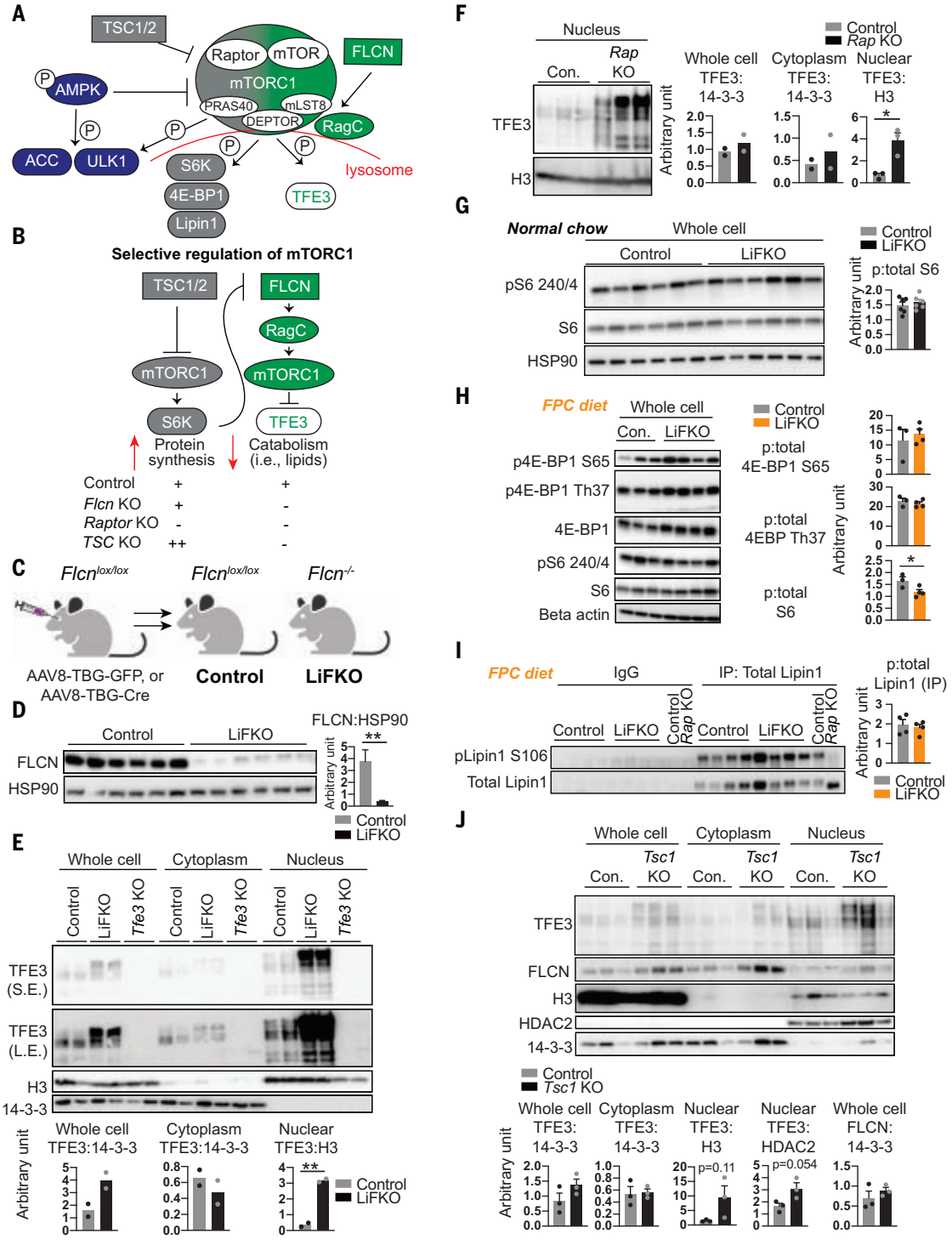


Fig. 1. Liver FLCN selectively promotes mTORC1-mediated cytoplasmic sequestration of TFE3, without affecting canonical mTORC1 signaling. (**A** and **B**) Model of selective mTORC1 regulation. P, phosphorylated; PRAS40, proline-rich AKT substrate of 40 kDa; DEPTOR, DEP domain-containing mTOR-interacting protein; mLST8, target of rapamycin complex subunit LST8. (**C**) Schematic of FLCN liver deletion to yield control and hepatocyte-specific *Flcn*-null (LiFKO) mice. (**D**) FLCN protein expression in livers of control and LiFKO mice fed normal chow and euthanized after overnight fast followed by 4 hours of refeeding. HSP90, heat shock protein 90. (**E**) TFE3 protein expression in the indicated subcellular fractions from livers of control, LiFKO, and *Tfe3* knockout mice fed normal chow and euthanized ad lib. S.E., short exposure; L.E., long exposure; H3, histone H3. (**F**) TFE3 protein expression in the nucleus from livers of control (Con.) and *Raptor* liver-KO mice, euthanized after overnight fasting and 4 hours of refeeding. See fig. S1A for whole cell and cytoplasmic fractions. p:total, fraction of phosphorylated protein to total protein. (**G** and **H**) Phosphorylation of mTORC1 targets in control and LiFKO mice fed either (G) normal chow or (H) 7 or 8 days of FPC diet (TD190142) and sugar water, and euthanized after overnight fast and ~4 hours of refeeding. pS6 240/4, phospho-ribosomal protein S6 ($Ser^{240/244}$); S6, ribosomal protein S6; p4E-BP1 S65, phospho-4E-BP1 (Ser^{65}); p4E-BP1 Th37, phospho-4E-BP1 (Thr^{37}). (**I**) Phosphorylation of Lipin1, assayed by immunoprecipitation (IP) with anti-Lipin1 followed by immunoblotting for phospho-Lipin1, in liver lysates of control and LiFKO mice fed 9 days of FPC diet (TD190142) and sugar water and euthanized at 10 p.m. ad lib. IgG, immunoglobulin G. (**J**) TFE3 protein expression in the indicated subcellular fractions from control and liver *Tsc1* KO mice fed normal chow and euthanized ad lib. *$P < 0.05$, **$P < 0.01$; Student's two-tailed *t* test. Data are depicted as mean ± SEM.

diet [the fructose, palmitate, and cholesterol (FPC) diet] (Fig. 1H). Phosphorylation of S6 was unchanged in mice fasted overnight and then refed normal chow (Fig. 1G), and only mildly decreased in mice fed an FPC diet (Fig. 1H). Phosphorylation of 4E-BP1 was unchanged (Fig. 1H). Phosphorylation of Lipin1, an mTORC1 substrate reported to inhibit SREBP-1c in the liver (*8*), was also unaffected in LiFKO livers (Fig. 1I). Together, these data indicate that hepatic FLCN activates mTORC1 selectively to regulate TFE3 localization, which is known to be regulated by mTORC1-mediated phosphorylation, with little impact on canonical mTORC1 substrates (Fig. 1, A and B).

FLCN binds to and inhibits or activates adenosine monophosphate–activated protein kinase (AMPK) in various cell types (*24*–*26*). We observed a small increase in phospho-AMPK (active AMPK) in livers from chow-fed LiFKO mice (fig. S1B) but no change in mice fed a NAFLD-inducing diet (fig. S1C). Phosphorylation of AMPK substrate acetyl–coenzyme A (acetyl-CoA) carboxylase (ACC) was unchanged in animals fed normal chow but increased in mice fed an FPC diet, whereas phosphorylation of unc-51–like autophagy kinase 1 (ULK1) was largely unchanged in both conditions (fig. S1, B and C). Loss of FLCN thus appears to have variable but mild effects on AMPK signaling in the liver.

We next tested the impact of *Tsc1* deletion on TFE3 nuclear localization. The TSC complex suppresses canonical mTORC1 activity (*27*). TFE3 was localized to the nucleus in livers depleted of *Tsc1* (Fig. 1J), despite activation of canonical mTORC1 signaling (fig. S1D). We also observed selective stabilization of the slower migrating TFE3 band, a signature of FLCN suppression (Fig. 1J). Canonical mTORC1 signaling (e.g., through S6K) therefore appears to exert negative feedback on the FLCN:TFE3 axis of mTORC1 signaling, such that chronic activation of the canonical arm, achieved by *Tsc1* deletion, might suppress FLCN activity (Fig. 1, A and B). FLCN protein abundance was unchanged in livers lacking *Tsc1* (Fig. 1J), indicating that inhibition of FLCN occurs after translation. Thus, activation of TFE3 in the liver can be achieved, seemingly paradoxically, by either inhibition of (noncanonical) mTORC1 or activation of (canonical) mTORC1. The presence of this feedback might explain the observations that both activation and inhibition of mTORC1 protect against steatosis (*8*, *11*). In this model (Fig. 1, A and B), loss of *Flcn* should protect against NAFLD, and that protection should be dependent on TFE3 activity.

Loss of FLCN in the liver protects against NAFLD

LiFKO mice fed standard rodent chow appeared grossly normal, gained body weight at similar rates to control mice (Fig. 2A), and exhibited normal liver histology, as shown by hematoxylin and eosin (H&E) staining (Fig. 2B, top, and fig. S2A). When fed an Amylin liver NASH (AMLN) diet—a NAFLD-inducing diet high in trans fat, fructose, and cholesterol (*28*)—control mice gained weight (Fig. 2C) and developed severe liver steatosis, as shown by H&E staining (Fig. 2B, bottom), liver triglyceride quantification (Fig. 2D), and blinded histological evaluation of H&E slides (Fig. 2E). In contrast, LiFKO mice were mostly protected from both body weight gain and liver steatosis (Fig. 2, B to E). LiFKO mice also revealed lower levels of plasma triglycerides (fig. S2B) and nonesterified fatty acids (fig. S2C) and were protected from diet-induced elevations in total cholesterol, both HDL and non-HDL (fig. S2, D to F). Similar protection from weight gain and steatosis was observed in mice fed a Gubra AMLN (GAN) diet, which is a modified AMLN diet without trans fat that has replaced the AMLN diet since the ban on trans fats (*29*) (fig. S2G). Loss of *Flcn* is thus protective against NAFLD in these models.

One possible explanation for the observed protection against steatosis in LiFKO mice was the protection against weight gain and concomitant reductions in plasma insulin and homeostatic model assessment for insulin resistance (HOMA-IR) (fig. S2, H to J). The mechanism for the protection against weight gain in LiFKO mice on an AMLN diet is not clear, as comprehensive lab animal monitoring system (CLAMS) studies with control and LiFKO mice fed a GAN diet revealed no significant differences in food consumption, water consumption, ambulatory or locomotor activity, or energy expenditure (fig. S3). We thus sought a diet where body weight gains were equivalent between genotypes. Control and LiFKO mice were subjected to an FPC diet regimen, consisting of high fat, sucrose, and cholesterol with reduced vitamin E and choline, in conjunction with fructose and glucose in the drinking water—often called the "Big Mac and Coke" diet (*30*). With this FPC regimen, in which diminished choline intake can suppress weight gain (*30*, *31*), there were no significant differences in body weight between the genotypes (Fig. 2H and fig. S4A), despite an apparent mild reduction in energy expenditure (fig. S4B). Plasma triglycerides, nonesterified fatty acids, insulin, glucose, and HOMA-IR were also unaffected (fig. S4, C to G). However, as with the AMLN diet, LiFKO mice were almost entirely protected against NAFLD induced by the FPC regimen (Fig. 2, G to J). The beneficial effects of *Flcn* deletion on the liver thus appear to be independent of body weight.

NAFLD protection through loss of FLCN requires TFE3

Loss of *Flcn* induces TFE3 translocation to the nucleus (Fig. 1E). Notably, codeletion of *Flcn* and *Tfe3* [double knockout (DKO) mice] completely prevented the protection against NAFLD seen with *Flcn* deletion alone, in both the AMLN and FPC diets (Fig. 2, B, D, E, G, I, and J). *Tfe3* deletion alone had little impact compared with control mice (Fig. 2, I to J). In the AMLN diet cohort, deletion of *Tfe3* also reversed the reductions in body weight seen with *Flcn* deletion (Fig. 2C), whereas no impact on body weight was seen in the FPC diet cohort (Fig. 2H). The protection against NAFLD in *Flcn* liver-null mice therefore requires TFE3. Together, these data demonstrate that loss of liver *Flcn* strongly protects against the development of NAFLD, in at least two NAFLD-inducing dietary conditions; that the protection is independent of effects on body weight; and that the protection is mediated by TFE3, most likely through TFE3 dephosphorylation, nuclear translocation, and transcriptional activation of an antisteatotic program.

Loss of FLCN in the liver activates pathways of lipid catabolism

In many cell types, TFE3 activates the expression of genes that drive lysosomal biogenesis (*15*, *32*), an important step in the breakdown of lipids through lipophagy (*33*), and the expression of PGC-1α and PGC-1β (*14*), drivers of mitochondrial biogenesis, fatty acid oxidation, electron transport chain, and tricarboxylic acid (TCA) cycle genes. RNA sequencing (RNA-seq) analysis (Fig. 3, A to C) and protein immunoblotting (Fig. 3D) showed that both of these TFE3-mediated gene programs were activated in livers from LiFKO mice, whether fed normal chow or a NAFLD-inducing diet. Moreover, the induction of both programs was abrogated in DKO livers (Fig. 3, B and D), demonstrating their dependence on TFE3. PGC-1α mRNA (*Ppargc1a*) abundance was increased in LiFKO livers in a TFE3-dependent manner (Fig. 3E), and TFE3 occupancy on the chromatin at *Ppargc1a*, as well as lysosomal genes [e.g., cathepsin z (*Ctsz*)], was increased in LiFKO livers, coinciding with PolII occupancy and epigenetic markers of transcriptionally active chromatin (Fig. 3F). In line with these data, fatty acid oxidation, measured by conversion of radiolabeled palmitate to water, was increased in hepatocytes isolated from LiFKO mice (Fig. 3G). Gluconeogenic targets of PGC-1α (*34*) did not show increased expression in *Flcn* KO livers (fig. S5A), indicating that deletion of *Flcn* activates PGC-1α–driven oxidation programs specifically, without adversely affecting gluconeogenesis. Consistent with this observation, glucose sensitivity was unchanged in mice fed a long-term AMLN diet (fig. S5B) or a short-term GAN diet (fig. S5C) and, if anything, improved in LiFKO mice fed normal chow (fig. S5B). Together, these data indicate that *Flcn* deletion, by releasing TFE3 to the nucleus, allows TFE3 to activate a coordinated program of lipid catabolism, thereby potentially limiting liver steatosis. However, the overall increase in systemic fatty acid oxidation seemed limited, as measured by β-hydroxybutyrate levels in plasma (fig. S6A) and whole-body respiratory exchange ratio values (fig. S6B). In light of the strong protection against NAFLD observed in LiFKO mice, these observations suggested that loss of *Flcn* might promote other antisteatotic programs.

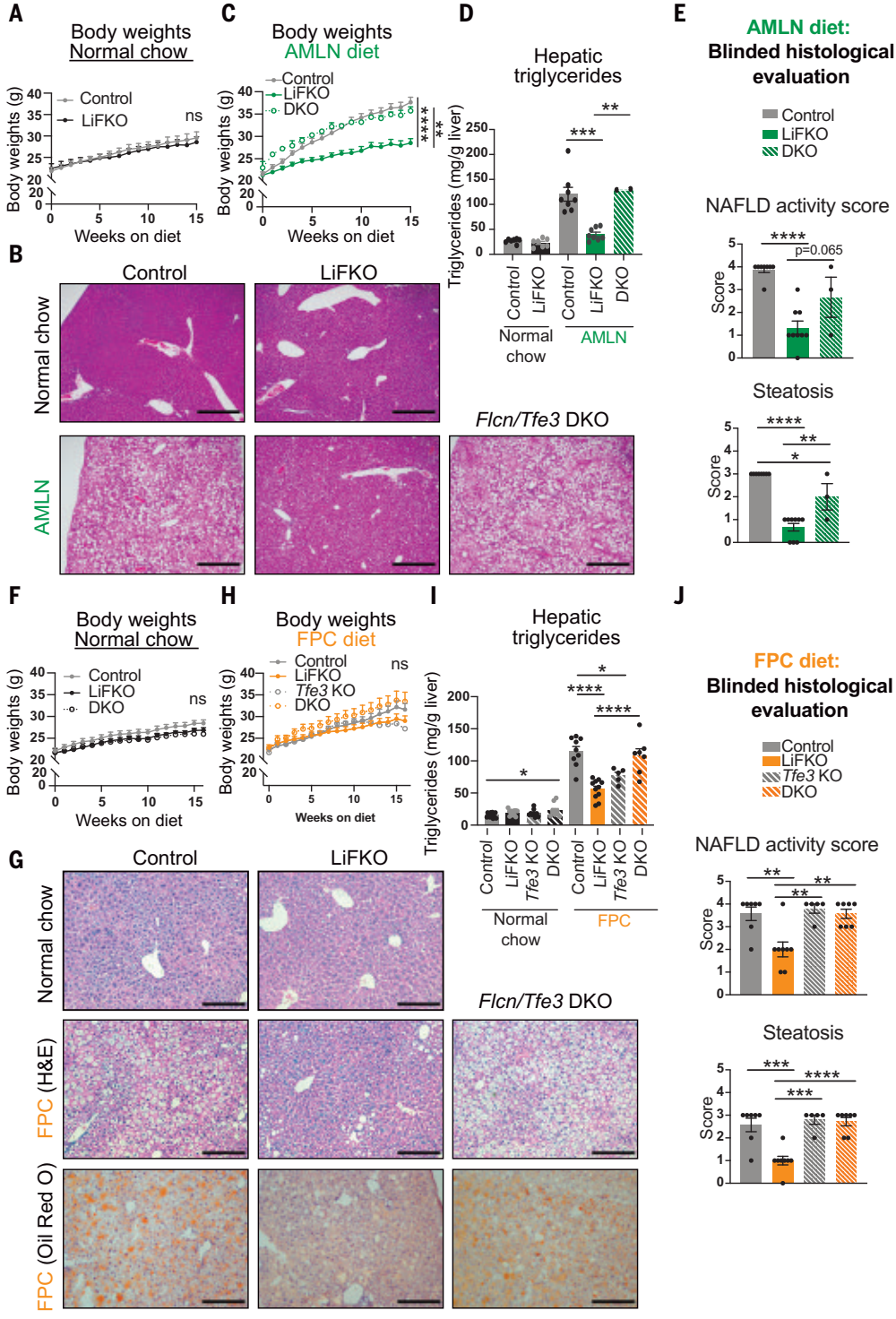


Fig. 2. Loss of FLCN in the liver potently protects against NAFLD, and the protection requires TFE3. (**A** to **E**) Control, LiFKO, and DKO mice were fed a normal chow (n = 7) or AMLN diet (n = 3 to 9) for 17 to 18.5 weeks and euthanized after a 4- to 6-hour fast. (A) Body weights on normal chow. (B) Representative images of liver H&E staining. Scale bars, 500 μm. (C) Body weights on AMLN diet. (D) Quantification of liver triglycerides. (E) Blinded histological evaluation of liver H&E slides. (**F** to **J**) Control, LiFKO, *Tfe3* KO, and DKO mice were fed normal chow (n = 10 to 13) or FPC diet regimen (TD160785 with sugar water; n = 5 to 11) for 16 weeks and euthanized after removing their food for 4 to 6 hours. (F) Body weights on normal chow. (G) Representative images of liver H&E staining. Scale bars, 200 μm. (H) Body weights on FPC diet. (I) Quantification of hepatic liver triglycerides. (J) Blinded histological evaluation of liver H&E slides. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001; ns, not significant. Statistical values are indicated for the final body weight measurement. Statistical analysis was done with two-way repeated measures or mixed effects analysis of variance (ANOVA) with multiple comparisons test in (A), (C), (F), and (H). Student's two-tailed t test was used for normal chow controls for the AMLN diet experiment. One-way ANOVA with Tukey's multiple comparisons test was used otherwise. Data are depicted as mean ± SEM.

Loss of FLCN in the liver suppresses de novo lipogenesis

We had expected that *Flcn* deletion would promote lysosomal and mitochondrial biogenesis. However, RNA-seq and subsequent quantitative polymerase chain reaction (qPCR) and Western blotting revealed that it also caused suppression in LiFKO livers, in both normal chow and NAFLD-inducing diet conditions, of a broad program of de novo lipogenesis gene and protein expression, including adenosine triphosphate citrate lyase (*Acly*), fatty acid synthase (*Fasn*), acyl-CoA synthetase short chain family member 2 (*Acss2*), stearoyl-CoA desaturase-1 (*Scd1*), sterol regulatory element–binding transcription factor 1 (*Srebf1*), and specifically the *Srebp1c* isoform of *Srebf1*, which encodes for SREBP-1c, the predominant driver of de novo lipogenesis in the liver (Fig. 4, A to D, and fig. S7A). Most of these observed gene and protein expression changes were reversed in DKO livers, and *Tfe3* deletion alone had little impact compared with control mice (Fig. 4, A to D, and fig. S7A). We conclude that loss of *Flcn* in the liver broadly suppresses the de novo lipogenesis gene program and appears to do so in a manner dependent on TFE3.

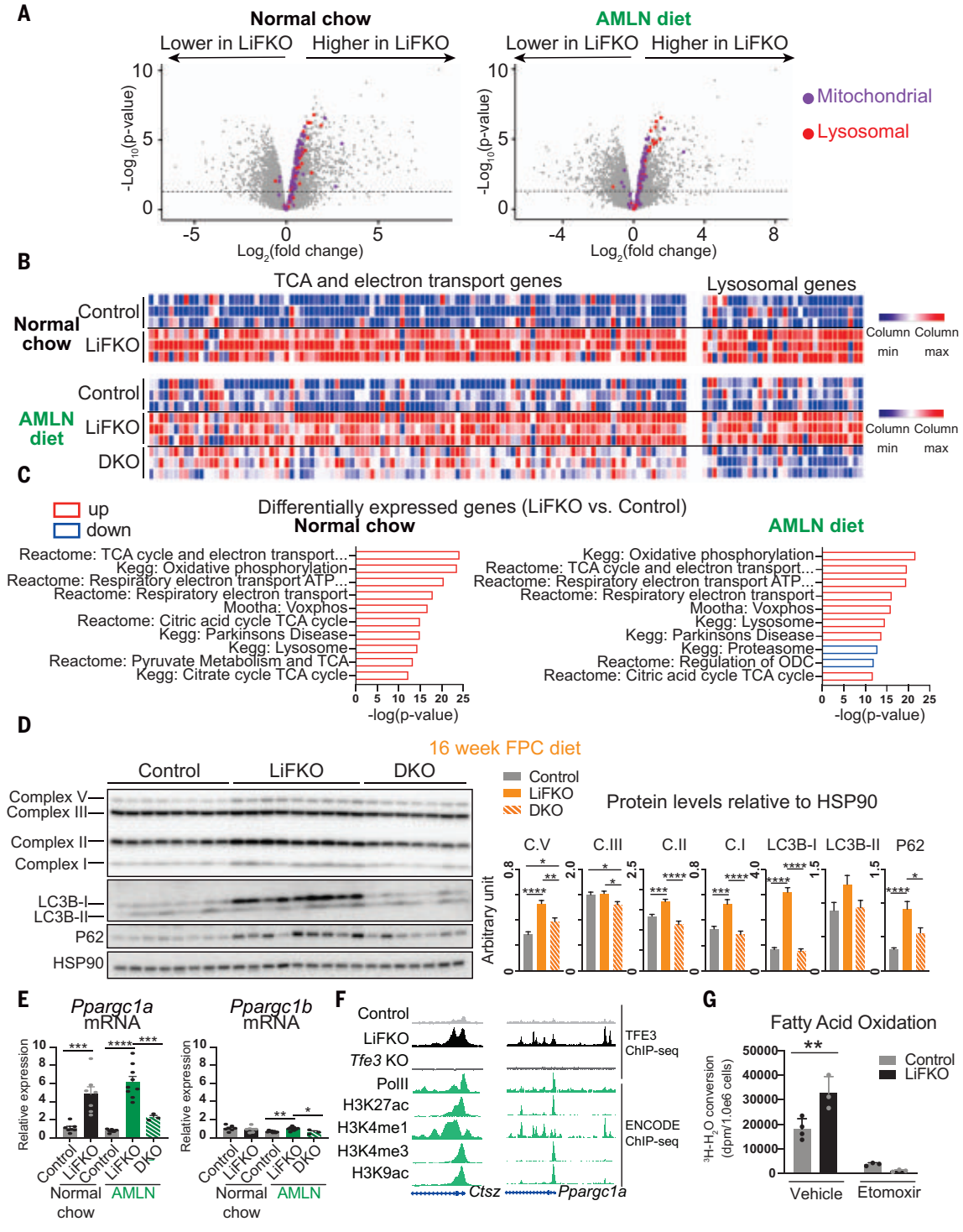


Fig. 3. Loss of FLCN in the liver activates pathways of lipid catabolism. (**A** to **C**) RNA-seq was performed on the livers of control, LiFKO, and DKO mice (n = 3) on normal chow or AMLN diet (as described in Fig. 2). (A) $-\text{Log}_{10}$(adjusted p-value) versus $\log_2$(LiFKO/control fold change) volcano plot. Dotted line: P = 0.05. (B) Heatmap of normalized expression of mitochondrial and lysosomal gene sets. (C) Top 10 differentially expressed gene sets between LiFKO and control livers. Red indicates pathways up-regulated in LiFKO, blue indicates pathways down-regulated in LiFKO. (**D**) Protein expression of components of mitochondrial electron transport chain complexes (C.I, C.II, C.III, and C.V), LC3B-I and LCSB-II, p62, and HSP90, in livers of control, LiFKO, and DKO mice fed an FPC diet regimen (TD160785 with sugar water) for 16 weeks. (**E**) mRNA expression of *Ppargc1a* and *Ppargc1b* in livers of control, LiFKO, and DKO mice fed a normal chow (n = 7) or AMLN diet (n = 3 to 9) for 17 to 18.5 weeks. (**F**) Genome browser tracks of the *Ctsz* and *Ppargc1a* promoters. TFE3 ChIP-seq was performed on livers of mice fed normal chow (n = 2 to 4). Depicted are tracks from one representative sample per genotype. Green indicates publicly available ENCODE liver ChIP-seq datasets. (**G**) Hepatocytes were isolated from three control and LiFKO mice and pooled. Fatty acid oxidation was measured from three or four wells of each genotype by incubating them with ^{3}H-palmitate for 120 min and either vehicle or 100 μM etomoxir (an inhibitor of fatty acid oxidation). Conversion of ^{3}H-palmitate to ^3H_2O was measured by scintillation counting and normalized to cell count. dpm, disintegrations per minute. One-way ANOVA with Tukey's multiple comparisons test (for three groups) or Student's two-tailed t test (for two groups) was used. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Data are depicted as mean ± SEM.

To measure DNL directly, we force-fed mice ^{13}C-fructose, a rich source of carbons for de novo lipogenesis in the liver, and quantified ^{13}C incorporation into liver fatty acids. De novo lipogenesis was decreased by about one-half in LiFKO mice fed a short-term AMLN diet, in the absence of changes in body weight (Fig. 4E). DKO animals had de novo lipogenesis slightly greater than that of LiFKO alone, but the effect was not statistically significant (Fig. 4E). Similar results were observed in mice injected with deuterium oxide, followed by measuring incorporation of ^{2}H into liver fatty acids, an alternate method of measuring de novo lipogenesis (Fig. 4F). Mild suppression

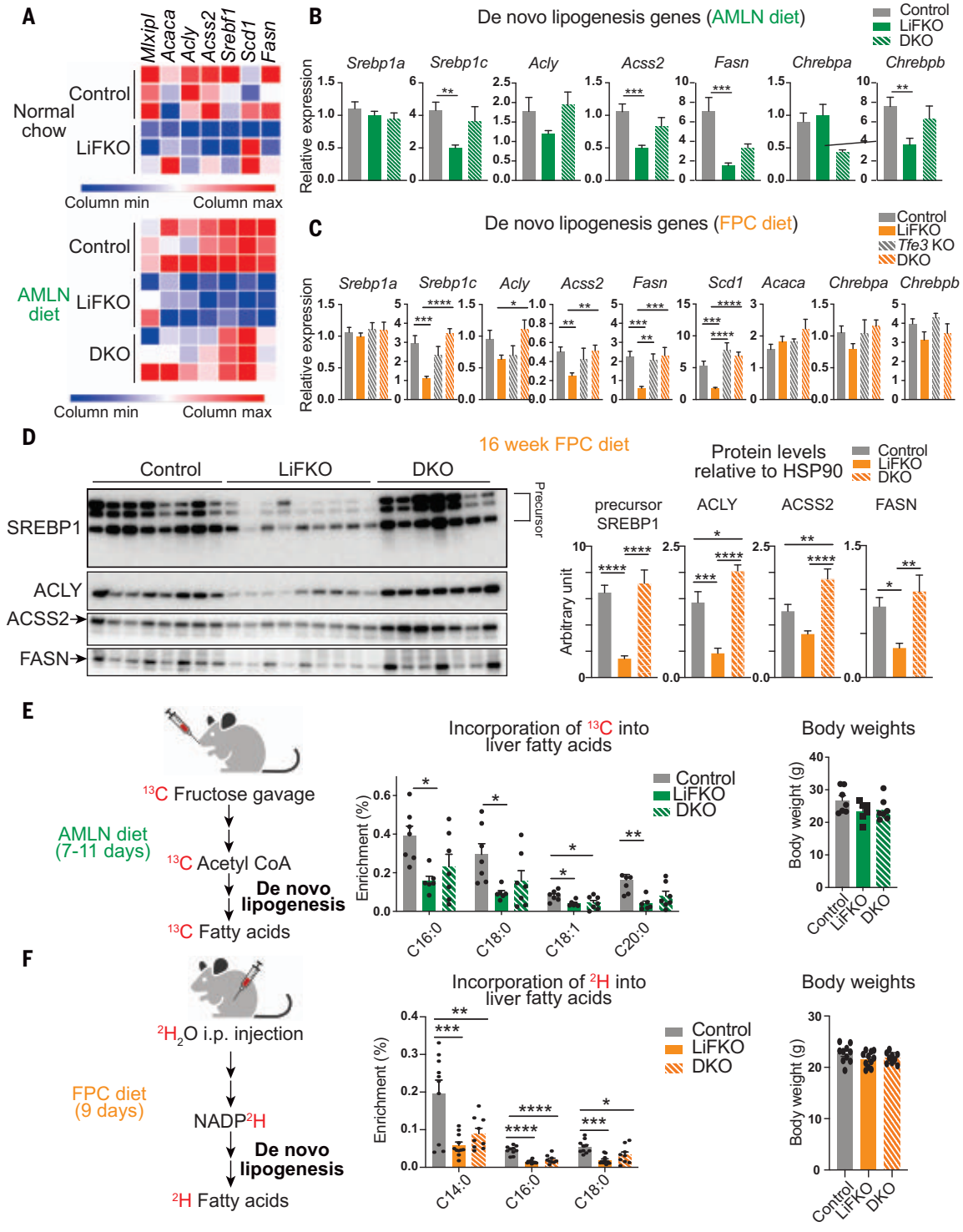


Fig. 4. Loss of FLCN in the liver suppresses de novo lipogenesis. (**A**) Heatmap of normalized expression values of de novo lipogenesis genes from RNA-seq described in Fig. 3. (**B** and **C**) mRNA expression of de novo lipogenesis genes in livers of control, LiFKO, *Tfe3* KO, and DKO mice fed (B) an AMLN diet (n = 3 to 9) for 17 to 18.5 weeks or (C) an FPC diet regimen (TD160785 with sugar water; n = 5 to 11) for 16 weeks. (**D**) Protein expression of SREBP1, ACLY, ACSS2, and FASN in livers from control, LiFKO, and DKO mice fed an FPC diet regimen (TD160785 with sugar water) for 16 weeks. HSP90 from Fig. 3D was used as loading control. Mice were euthanized after a 4- to 6-hour fast, and thus only the precursor form of SREBP1 was detected. (**E**) Control, LiFKO, and DKO mice (n = 6 or 7) were fed an AMLN diet for 7 to 11 days, fasted from 9 a.m. to 7 p.m., refed for 2 hours, and force-fed a bolus of ^{13}C-fructose and ^{12}C-glucose. The mice were fed overnight and killed the next morning. LC-MS was performed to examine the amount of ^{13}C label incorporation into hepatic fatty acids. (**F**) Control, LiFKO, and DKO mice (n = 9 to 11) were fed an FPC diet regimen (TD190142 with sugar water) for 9 days and then injected intraperitoneally (i.p.) with deuterium oxide ($^{2}H_2O$) at ~7 p.m. Five hours later, the mice were killed, and their livers harvested. LC-MS was performed to examine the amount of deuterium label incorporation into hepatic fatty acids. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001; one-way ANOVA with Tukey's multiple comparisons test. Data are depicted as mean ± SEM.

of DNL was also evident in LiFKO mice fed normal chow (fig. S7B). Together, these data demonstrate that the FLCN:mTORC1:TFE3 axis in the liver both induces a catabolic (lysosomal and oxidative) gene program and simultaneously suppresses anabolic de novo lipogenesis.

TFE3 acts downstream of LXR to suppress de novo lipogenesis

Transcription of de novo lipogenesis genes is largely mediated by the transcription factor SREBP-1c (*35*). SREBP-1c activity is regulated at both the protein and mRNA levels. Thus, TFE3 might suppress de novo lipogenesis genes by inhibiting the SREBP-1c pathway at one or more points of regulation. The transcription factor liver X receptor (LXR) promotes transcription of SREBP-1c (*35*). We therefore tested whether TFE3 suppressed LXR. RNA-seq data revealed that some LXR target genes were suppressed in LiFKO livers but others were not (fig. S8A). The expression of liver X receptor alpha (LXR alpha; *Lxra*) and LXR beta (*Lxrb*) was unchanged in LiFKO livers from mice under various dietary conditions (fig. S8B). We next treated LiFKO mice with the LXR agonist T0901317 (T09) in conjunction with long-term FPC diet (fig. S8, C and D). T09 significantly exacerbated liver steatosis in control mice (*36*) (fig. S8E). Deletion of liver *Flcn*, however, suppressed liver steatosis, reversing the increase in liver triglyceride content promoted by T09 (fig. S8E). Analyses of liver de novo lipogenesis genes and other LXR target gene expression revealed similar patterns, whereby de novo lipogenesis genes were uniformly suppressed in LiFKO mice even in the presence of LXR agonism (fig. S8F). These data demonstrate that loss of *Flcn* and activation of TFE3 act downstream of LXR to suppress DNL.

TFE3 suppresses SREBP-1c proteolytic processing and activation

We tested whether *Flcn* deletion inhibited SREBP-1c posttranslational regulation. SREBP-1c exists as an inactive precursor (pSREBP-1c) in the endoplasmic reticulum (ER) and is proteolytically cleaved to liberate its basic helix–loop–helix leucine zipper (bHLH-Zip) domain (nSREBP-1c) to translocate to the nucleus and transcriptionally activate target genes, including de novo lipogenesis genes (*37*). The proteins insulin-induced gene 1 (INSIG1) and 2 (INSIG2) suppress processing of pSREBP-1c (*38*). To assay SREBP-1c processing, we killed control and LiFKO mice at night during feeding, when SREBP-1c proteolytic processing is active (*39*, *40*). After 9 days of FPC diet, chosen to model an early steatotic state with no differences in body weight (Fig. 5A), we observed an increased abundance of the liver-specific *Insig2a* isoform (*11*, *41*) of *Insig2*, as well as INSIG2 protein, in the livers of LiFKO mice (Fig. 5, B and C). The accumulation of *Insig2* mRNA and of INSIG2 protein required TFE3, as it was absent in the DKO mice (Fig. 5, B and C). Chromatin immunoprecipitation sequencing (ChIP-seq) experiments showed chromatin occupancy by TFE3 in the *Insig2* regulatory region, which was further increased in livers from LiFKO mice, consistent with direct regulation of *Insig2* mRNA expression by TFE3 (Fig. 5D). In line with the established role of INSIG2 in suppressing SREBP-1c processing, abundance of the fully processed nSREBP-1c form (designated "N" in Fig. 5C) was decreased in LiFKO livers, reflecting the large amounts of INSIG2 protein (Fig. 5C). The suppression of nSREBP-1c was less in DKO livers (Fig. 5C). Thus, FLCN appears to suppress expression of de novo lipogenesis genes in large part by suppressing proteolytic processing and activation of SREBP-1c (Fig. 5E).

To test this notion directly, we introduced, using AAV8 delivery, hemagglutinin (HA)–tagged constitutively processed nSREBP-1c to control and LiFKO mice (Fig. 5F). After 9 days, mice displayed similar body weights (Fig. 5G) and equal amounts of HA-nSREBP-1c protein in the liver (Fig. 5H). AAV8-nSREBP-1c nearly quadrupled triglyceride content in control mice (Fig. 5I) and increased transcription of de novo lipogenesis genes (Fig. 5J) (*42*). Hepatic triglycerides were elevated to a nearly equal extent in LiFKO and control mice infected with AAV8-nSREBP-1c (Fig. 5I), indicating that loss of FLCN did not suppress de novo lipogenesis in the presence of constitutive nuclear nSREBP-1c. Taken together, these data are consistent with loss of liver FLCN leading to TFE3 activation, increased expression of INSIG2, and suppression of SREBP-1c processing and activation, ultimately leading to suppression of de novo lipogenesis and hepatic steatosis.

TFE3 synergistically occupies chromatin in close proximity to SREBP-1c genome-wide

Although loss of *Flcn* did not suppress nSREBP-1c–induced accumulation of hepatic triglycerides (Fig. 5I), it did still mildly suppress expression of nSREBP-1c–induced DNL genes (Fig. 5J). This suggested that, in addition to inhibition of SREBP-1c processing through expression of INSIG2, TFE3 may also have a direct effect on processed, nuclear SREBP-1c. We therefore used ChIP-seq experiments to test whether TFE3 and SREBP-1c occupied similar, or the same, sites on chromatin. TFE3-bound chromatin regions were enriched for the TFE family motif, and pathway analyses revealed enrichment of "vesicle-mediated transport" and "membrane trafficking" (fig. S9), consistent with the known roles of the TFE family (*32*, *43*). The pathways of "metabolism" and "metabolism of lipids and lipoproteins" were consistently the two most enriched pathways among the TFE3 cistromes (fig. S9). Sites known to be occupied by SREBP-1c (*42*) overlapped with TFE3 binding, which was increased in the absence of FLCN, in mice fed normal chow or an FPC diet (fig. S10A). We conclude that TFE3 binds genome-wide at genes of lipid homeostasis, that binding is increased by loss of FLCN, and that these TFE3 binding sites overlap with SREBP-1c binding sites.

TFE3, like SREBP-1c, is a bHLH-Zip transcription factor that binds to DNA containing E-box motifs (*44*, *45*). This raised the possibility that TFE3 and SREBP-1c may compete for access to chromatin. However, binding by HA-nSREBP-1c to its cognate chromatin sites was no different between control and LiFKO mice (fig. S10B), indicating that there is no competitive removal of SREBP-1c from chromatin by TFE3. In the converse experiment, genome-wide TFE3 binding in these same livers increased in the presence of exogenous nSREBP-1c (fig. S10C). Unbiased motif analyses of TFE3 binding sites that increased in the presence of AAV8-HA-nSREBP-1c identified the SREBP-responsive element as enriched (fig. S10D). Thus, expression of nSREBP-1c promotes binding of TFE3 at, or near, SREBP-1c binding sites. Closer evaluation of TFE3 occupancy specifically at sites near de novo lipogenesis gene transcription start sites revealed a pattern, seen only in the absence of FLCN, whereby TFE3 bound chromatin closely downstream of SREBP-1c binding (fig. S10, E and F).

These data indicate that loss of FLCN induces changes in both TFE3 binding pattern and binding intensity and that TFE3 binding localizes near SREBP-1c binding genome-wide. However, these factors do not compete with each other for binding to chromatin; instead, the presence of SREBP-1c synergistically promotes binding of TFE3.

Loss of FLCN in mouse liver prevents and reverses NASH

NAFLD can progress to NASH, a predictor of poor clinical outcomes in humans (*46*). The choline-deficient amino acid–defined and high fat (CDAA-HF) diet is a rapid NASH-inducing diet that recapitulates some aspects of human physiology (*47*, *48*). Unlike mice on other NASH-inducing diets, mice on the CDAA-HF diet do not lose weight (Fig. 6A). Control mice fed the CDAA-HF diet showed liver steatosis and NASH, as demonstrated by tripling of Sirius Red staining for fibrosis (Fig. 6, B and C); greater than 30-fold increase in collagen type I alpha 1 (*Col1a1*) gene expression (Fig. 6D), a marker for NASH (*49*); and increased expression of a range of genes involved in fibrosis and inflammation (Fig. 6E). In contrast, LiFKO mice were largely protected from NASH, as seen in H&E and Sirius Red staining (Fig. 6, B and C), suppression of many genetic

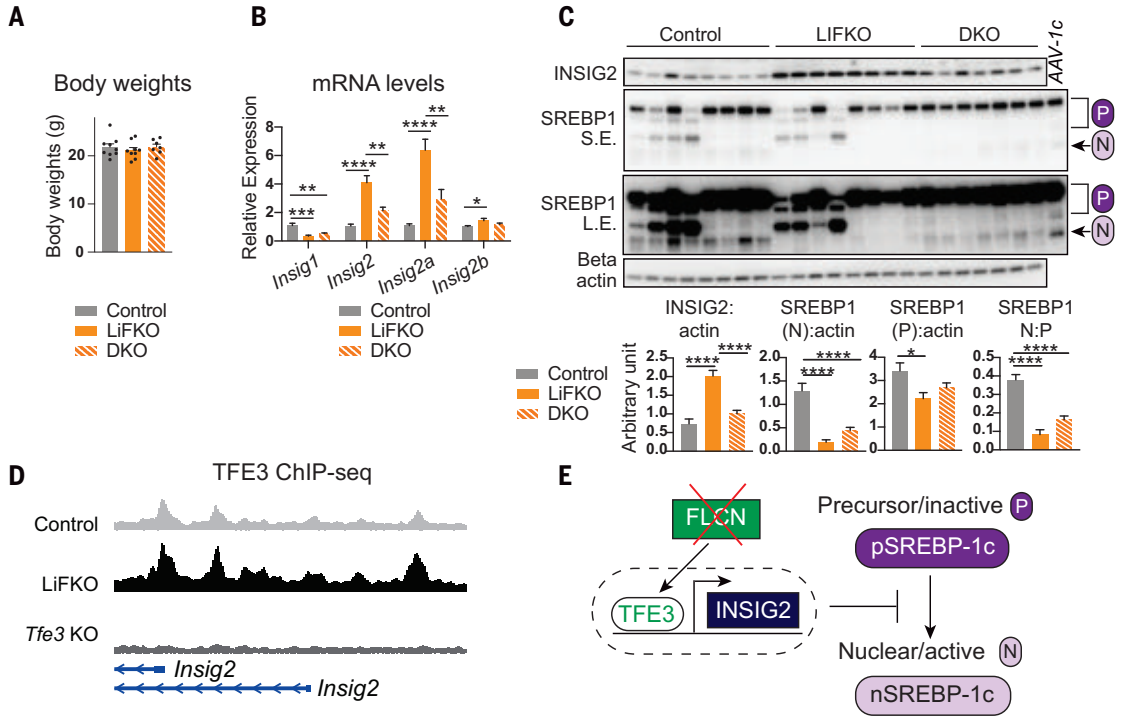


Fig. 5. TFE3 suppresses SREBP-1c proteolytic processing and activation. (**A** to **D**) Control, LiFKO, and DKO mice were fed an FPC diet regimen (TD190142 with sugar water) for 9 days and euthanized at 10 p.m. ad lib (n = 7 to 9). One-way ANOVA with Tukey's multiple comparisons test was used. (A) Body weights. (B) mRNA expression of DNL genes. (C) Liver protein expression of SREBP-1, INSIG2, and beta actin. AAV-1c, positive control from liver injected with AAV8 expressing constitutively nuclear SREBP-1c; P, precursor form of SREBP-1; N, nuclear (processed) form of SREBP-1. (D) Genome browser tracks of the *Insig2* promoter. Depicted are TFE3 ChIP-seq tracks (described in Fig. 3F) from one representative sample per genotype. (**E**) Schematic of FLCN:TFE3 regulation of SREBP-1c proteolytic processing. (**F** to **J**) *Flcn*$^{lox/lox}$ mice (n = 5 or 6) were injected with either "control virus" (AAV8-GFP or AAV8-Cre; to generate control or LiFKO mice) or AAV8-ApoE/AAT-HA-nSREBP-1c ("1c") and then maintained for 9 days on an FPC diet (TD190142 with sugar water). Student's two-tailed t test was used for analysis. (F) Experimental outline. (G) Body weights. (H) Liver protein expression of exogenous HA-tagged nuclear SREBP-1c (HA), total SREBP-1, INSIG2, and 14-3-3. (I) Hepatic triglyceride quantification. (J) Liver mRNA expression of indicated DNL genes. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Data are depicted as mean ± SEM.

markers of fibrosis (Fig. 6, D and E), and suppressed steatosis and fibrosis measured through blinded histological evaluation (fig. S11). Deletion of TFE3 in DKO mice largely abrogated the protection afforded by FLCN deletion in LiFKO mice, demonstrating the critical role of TFE3 in protection against NASH, as with NAFLD (Fig. 6, D and E). Fibrosis levels remained low in histological evaluations even in the absence of TFE3 (Fig. 6C and fig. S11), indicating that FLCN also regulates other pathways to protect against NASH. Similar reductions of NASH markers were observed in LiFKO mice fed either the AMLN diet (fig. S12, A and C) or FPC diet (fig. S12, B and D), although the induction of NASH in control mice fed these diets was milder than that seen with the CDAA-HF diet. Thus, loss of liver FLCN protects against the development of NASH and NAFLD in multiple mouse models.

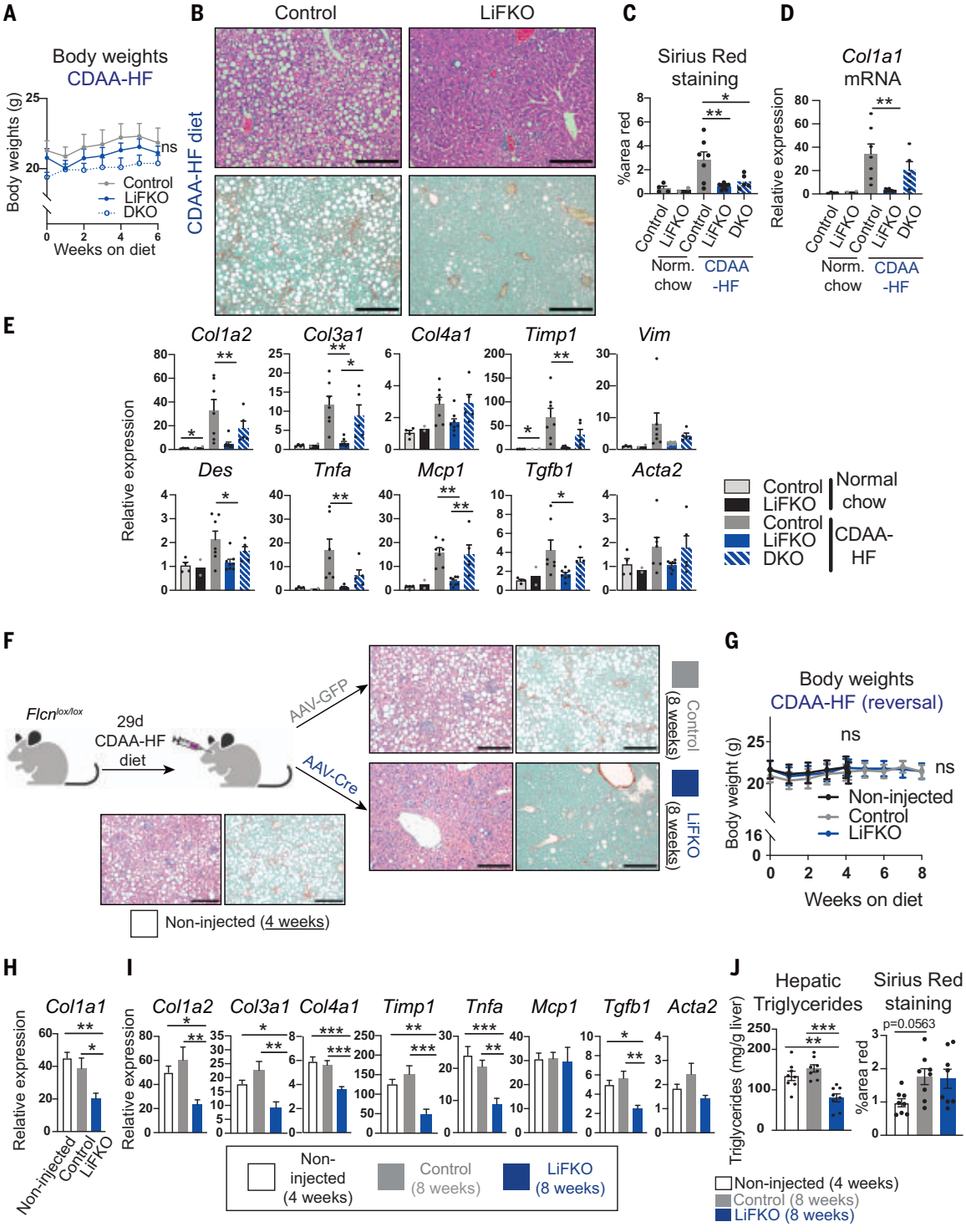


Fig. 6. Loss of FLCN in the liver prevents and reverses NASH. (**A** to **E**) Prevention of NASH. Control, LiFKO, and DKO mice were maintained on normal chow (n = 2 to 4) or CDAA-HF diet (n = 5 to 8) for 6 weeks and then euthanized after a 4- to 6-hour fast. (A) Weekly body weights. (B) Representative images of liver H&E and Sirius Red staining of mice fed CDAA-HF diet. (C) Quantification of Sirius Red positive staining. [(D) and (E)] Liver mRNA expression of the indicated fibrosis and inflammation markers. (**F** to **J**) Reversal of NASH. (F) Experimental outline. *Flcn*$^{lox/lox}$ mice were fed a CDAA-HF diet for 29 days to induce NASH, at which time "non-injected" mice were euthanized (n = 8). The remaining mice were injected with AAV8-TBG-GFP or AAV8-TBG-Cre to yield control and LiFKO mice (n = 8 in each group), and subsequently maintained for another 4 weeks on a CDAA-HF diet. Representative liver H&E and Sirius Red images depicted for each group. (G) Weekly body weights. [(H) and (I)] Liver mRNA expression of the indicated fibrosis and inflammation markers. (J) Quantification of hepatic liver triglycerides and Sirius Red positive staining. Two-way repeated measures ANOVA with multiple comparisons test were used in (A) and (G). One-way ANOVA with Tukey's multiple comparisons test was used to assess differences between CDAA-HF–fed control, LiFKO, and DKO mice [(C) to (E)] or noninjected, control, and LiFKO mice [(H) to (J)]. Student's two-tailed t test was used to compare control and LiFKO mice on normal chow. *P < 0.05, ** P < 0.01, *** P < 0.001, ****P < 0.0001. Scale bars, 200μm. Data are depicted as mean ± SEM.

A therapy intended to treat NASH and NAFLD would benefit from modifying established disease, in addition to preventing it. We therefore provided *Flcn*$^{lox/lox}$ mice with the CDAA-HF diet for 4 weeks to induce NASH (Fig. 6F) and then infected the mice with AAV8-GFP (green fluorescent protein) or AAV8-Cre to yield control and LiFKO mice, respectively, followed by another 4 weeks of CDAA-HF diet (Fig. 6F). There was no difference in body weights at either the 4- or 8-week time points (Fig. 6G). After 8 weeks, deletion of *Flcn* led to significant reversal of NASH and NAFLD when compared with both the 4-week pre-deletion mice and the 8-week control mice, as evidenced by reduced steatosis on H&E (Fig. 6F); reduced *Col1a1* expression (Fig. 6H); reduced expression of other measured genes of inflammation and fibrosis (Fig. 6I); reduced hepatic triglyceride accumulation (Fig. 6J); and reduced NAFLD activity scores, steatosis levels, and inflammation levels evaluated through blinded histological evaluation (fig. S11). Sirius Red staining and fibrosis scores were not reversed by *Flcn* deletion, perhaps reflecting the slower process of reversing collagen deposition (Fig. 6J and fig. S11). Taken together, these data demonstrate that suppression of liver FLCN prevents NASH as efficiently as it prevents NAFLD, that the suppression is largely dependent on TFE3, and that loss of FLCN can also reverse most characteristics of NAFLD and NASH.

Discussion

Suppression of FLCN in the liver protected mice from both NAFLD and NASH and helped reverse these processes if already established. The data thus reveal FLCN as a possible target for the treatment of NAFLD and NASH. Loss-of-heterozygosity or "second hit" somatic mutations in carriers of FLCN mutations can lead to renal cell cancer (*50*), raising a theoretical concern for possible development of cancer with any therapy that suppresses FLCN. However, FLCN suppression can be targeted specifically to the liver through, for example, liver-targeted nanoparticles or *N*-acetylgalactosamine (GalNAc)–modified small interfering RNAs (siRNAs) (*51*–*53*), and hepatocellular carcinoma has not been observed in carriers of FLCN mutations. Moreover, FLCN deletion is thought to induce renal cell cancer through induction of mTORC1 signaling (*54*, *55*), and we did not observe higher canonical mTORC1 activity in livers of LiFKO mice (Fig. 1, G and H).

Our data provide insight into the mechanisms by which mTORC1 regulates and coordinates lipid homeostasis in the liver. The mTORC1 pathway is sometimes depicted as receiving multiple inputs and integrating that information into a single on/off switch that then phosphorylates all of its targets. However, specificity of mTORC1 signal transduction in fact exists (*14*–*16*), and suppression of FLCN selectively suppresses mTORC1-mediated phosphorylation of TFE3, with minimal effect on other targets. This FLCN:mTORC1:TFE3 arm appears to be necessary for mTORC1-mediated lipid anabolism and steatosis in response to various diets. Moreover, cytoplasmic sequestration and suppression of TFE3 is the critical mediator of this anabolic signal.

This work provides at least a partial explanation for the seemingly contradictory data on the role of mTORC1 in the development of steatosis. We propose that the FLCN:mTORC1:TFE3 arm is dominant in the anabolic regulation of SREBP-1c and DNL. Consistent with this notion, deletion of *Flcn* itself is protective against NAFLD but has little impact on the canonical mTORC1:S6K arm. Overactivation of the mTORC1:S6K arm, achieved by deletion of *Tsc1*, may improve NAFLD (*10*–*12*) by activating a feedback loop that leads to suppression of the FLCN:mTORC1:TFE3 arm and to TFE3 nuclear translocation. Loss of *Tsc1* may thus promote protein anabolism through activation of the mTORC1:S6K arm but suppress lipid anabolism through feedback inhibition of the FLCN:mTORC1:TFE3 arm. These results are also consistent with the notion that the mTORC2 complex, which is known to promote DNL, does so directly rather than by suppression of TSC, adding another level of complexity to this system (*56*–*58*).

Our data provide mechanistic insight into the regulation of SREBP-1c and de novo lipogenesis, a critical component of NAFLD in humans (*59*, *60*). Releasing TFE3 to the nucleus induces INSIG2 to inhibit proteolytic processing of SREBP-1c and promotes TFE3 binding to chromatin near SREBP-1c to further suppress lipogenesis genes. This pathway of SREBP-1c regulation is independent of Lipin1 phosphorylation and of LXR signaling, indicating that Lipin1 is primarily targeted by the canonical arm of mTORC1, and that, under these conditions, Lipin1 is not necessary for efficient suppression of SREBP-1c and protection against NAFLD and NASH.

A potentially beneficial aspect of FLCN as a therapeutic target for NAFLD and NASH is its coordinated regulation of multiple programs involved in the progression to steatosis, that is, simultaneous induction of lipid consumption programs and suppression of lipid generation (de novo lipogenesis). This coordinated and synergistic effect likely drives the observed protection against NAFLD and NASH. Moreover, the coordinated effect on multiple arms of lipid handling may prevent the activation of feedback loops that may neutralize the benefits of targeting a single arm. For example, inhibitors against ACC have been tried in the clinic but were discontinued because of compensatory elevation in SREBP-1c activity and consequent hypertriglyceridemia (*61*). In light of these maladaptive responses to single-pathway inhibition, a strong argument has been made to identify upstream targets that coordinate control of many aspects of lipid metabolism in order to treat NASH successfully (*62*). We propose that FLCN is one such target.

This work has several limitations. Our studies were limited to genetic deletion of *Flcn*, but potential therapeutic approaches, such as liver-targeted GalNAc-modified siRNAs, are more likely to rely on incomplete suppression of FLCN. We also did not evaluate the impact of TFEB, a transcription factor with homology to TFE3 (*32*), the presence of which may in part explain why deletion of *Tfe3* does not reverse the consequence of *Flcn* deletion in some of our experiments. For example, differences between TFE3 and TFEB kinetics may explain why mice lacking both *Flcn* and *Tfe3* in the liver did not show significantly higher functional de novo lipogenesis relative to mice lacking only *Flcn* on a short-term NAFLD diet (Fig. 4, E and F) but did show higher de novo lipogenesis gene and protein expression on a long-term NAFLD diet (Fig. 4, B to D). For instance, it may be that the presence of TFEB compensates for TFE3 loss in the short term but not in the long term. It is also important to note that none of the diets used in our studies (AMLN, GAN, FPC, and CDAA-HF diets) perfectly recapitulate the characteristics of human NAFLD or NASH, and each diet was chosen to reflect some, but not all, characteristics (*28*, *30*, *47*, *48*). Finally, it is also important to note that, in the context of at least some of these diets (AMLN and GAN diets, but not FPC and CDAA-HF diets), LiFKO mice displayed reduced weight gain compared with control animals, indicating that deletion of *Flcn* in the liver may have additional systemic salutary effects.

Materials and methods summary

Most studies involved either adult homozygous *Flcn*$^{lox/lox}$ mice (*14*, *54*) or *Flcn*$^{lox/lox}$ mice also containing a loss-of-function whole-body *Tfe3* mutation (*14*, *45*), which were subsequently injected with ~1.5 × 10^{11} genome copies per mouse of AAV8-TBG-GFP or AAV8-TBG-Cre (Penn Vector Core, AV-8-PV0146 and AV-8-PV1091) to yield control mice (*Flcn*$^{lox/lox}$, *Tfe3*$^{+/Y}$ with GFP), LiFKO mice (*Flcn*$^{lox/lox}$, *Tfe3*$^{+/Y}$ with Cre), Tfe3 KO mice (*Flcn*$^{lox/lox}$, *Tfe3*$^{-/Y}$ with GFP), and DKO mice (*Flcn*$^{lox/lox}$, *Tfe3*$^{-/Y}$ with Cre). Mice were then fed various diets (normal chow, FPC diet regimen, AMLN diet, or GAN diet) and euthanized at various time points, and livers were harvested. For the AAV-nSREBP-1c rescue experiments, mice were additionally injected with 1.0 × 10^{11} genome copies per mouse of AAV8-ApoE/AAT-nSREBP-1c. We also used adult *Raptor*$^{lox/lox}$ mice injected with 1.0 × 10^{11} genome copies per mouse of either AAV8-TBG-GFP (control) or AAV8-TBG-Cre ("*Rap* KO" or

"*Raptor* KO") and euthanized 2 weeks later after an overnight fast followed by 4 hours of refeeding, as in (*12*) (Fig. 1F and fig. S1A). For the control versus *Tsc1* KO subcellular fractionation experiment, mouse liver samples from the laboratory of S. Biddinger were used. Eight- to nine-week-old female $Tsc1^{lox/lox}$ (control) mice or Albumin-Cre $Tsc1^{lox/lox}$ mice (*Tsc1* KO) were euthanized ad lib (*63*).

Radioimmunoprecipitation assay buffer lysis was used for protein isolation from whole liver tissues, followed by immunoblotting using standard wet-transfer methods; both commercially available and custom-made antibodies were used. To assay the cellular localization of TFE*3*, subcellular fractionation was done using whole cell lysis on mouse liver samples followed by centrifugation to yield nuclear and cytoplasmic fractions; immunoblotting was done using a commercially available anti-TFE*3* antibody. For phospho-Lipin1 analysis, total Lipin1 was immunoprecipitated using the custom-made anti-total Lipin1 antibody (*8*) and subsequently subject to immunoblotting using a custom-made anti-pLipin1 S106 antibody (*8*). Histology analysis was performed on liver tissue using standard paraformaldehyde fixation followed by ethanol dehydration and paraffin embedding; commercially available stains were then used to evaluate for triglyceride accumulation (H&E, Oil Red O) and fibrosis (Sirius Red). Liver triglycerides were quantified using a commercially available colorimetric reagent. Plasma metabolites (triglycerides, nonesterified fatty acids, cholesterol, insulin, glucose, and β-hydroxybutyrate) were all quantified using commercially available colorimetric reagents or an Axcel autoanalyzer, as described in the supplementary materials. Blood glucose was measured with a glucometer. For fatty acid oxidation measurements, primary hepatocytes were isolated, cultured overnight, and subjected to radioactive palmitate tracing as referenced (*42*). The cells were incubated with 125 μM ^{3}H-palmitate conjugated on bovine serum albumin and 1 mM carnitine for an additional 2 hours with or without 100 μM etomoxir. The media was delipidated, and $^{3}H_2O$ was measured by scintillation counting. Oral glucose tolerance tests was done on fasted mice gavaged with 2 g/kg D-glucose; blood glucose was then measured at serial time points. Whole-body metabolism and body composition were measured using the CLAMS and EchoMRI systems, respectively. RNA isolation and cDNA synthesis was performed with standard commercial extraction and synthesis methods. RNA-seq was done using GENEWIZ; data was analyzed with standard bioinformatic tools.

To measure de novo lipogenesis, mice were gavaged with a 1:1 mixture of ^{12}C D-glucose:^{13}C fructose (2 g/kg each) and euthanized; livers were subjected to liquid chromatography–mass spectrometry (LC-MS) to measure ^{13}C label incorporation into hepatic fatty acids. Alternatively, mice were injected intraperitoneally with 30 μl per gram of body weight of 99.9% D_2O; after euthanasia, livers were subjected to LC-MS to measure deuterium label incorporation into hepatic fatty acids. For the LXR agonist rescue experiment, mice were injected intraperitoneally with a commercially available LXR agonist (T0901317) or vehicle twice weekly, as described (*36*), while being fed an FPC diet regimen. ChIP-seq was performed by subjecting livers to sequential steps of nuclear isolation, fixation, sonication of chromatin, and immunoprecipitation with either anti-TFE*3* or anti-HA antibodies; crosslinks were then reversed, protein was degraded, and the remaining immunoprecipitated DNA was isolated. The ChIP input and immunoprecipitated DNA was then sequenced with commercially available library preparation and sequencing methods; data was analyzed with standard bioinformatic tools. Quantification and statistical analysis were done using GraphPad Prism, R, Python, and other standard methods. A full description of methods is provided in the supplementary materials.

REFERENCES AND NOTES

1. M. E. Rinella, Nonalcoholic fatty liver disease: A systematic review. *JAMA* **313**, 2263–2273 (2015). doi: 10.1001/jama.2015.5370; pmid: 26057287
2. S. Schuster, D. Cabrera, M. Arrese, A. E. Feldstein, Triggering and resolution of inflammation in NASH. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 349–364 (2018). doi: 10.1038/s41575-018-0009-6; pmid: 29740166
3. R. F. Schwabe, I. Tabas, U. B. Pajvani, Mechanisms of fibrosis development in nonalcoholic steatohepatitis. *Gastroenterology* **158**, 1913–1928 (2020). doi: 10.1053/j.gastro.2019.11.311; pmid: 32044315
4. A. C. Sheka *et al*., Nonalcoholic steatohepatitis: A review. *JAMA* **323**, 1175–1183 (2020). doi: 10.1001/jama.2020.2298; pmid: 32207804
5. Y. Kawano, D. E. Cohen, Mechanisms of hepatic triglyceride accumulation in non-alcoholic fatty liver disease. *J. Gastroenterol.* **48**, 434–441 (2013). doi: 10.1007/s00535-013-0758-5; pmid: 23397118
6. S. J. H. Ricoult, B. D. Manning, The multifaceted role of mTORC1 in the control of lipid metabolism. *EMBO Rep.* **14**, 242–251 (2013). doi: 10.1038/embor.2013.5; pmid: 23399656
7. C. C. Dibble, B. D. Manning, Signal integration by mTORC1 coordinates nutrient input with biosynthetic output. *Nat. Cell Biol.* **15**, 555–564 (2013). doi: 10.1038/ncb2763; pmid: 23728461
8. T. R. Peterson *et al*., mTOR complex 1 regulates lipin 1 localization to control the SREBP pathway. *Cell* **146**, 408–420 (2011). doi: 10.1016/j.cell.2011.06.034; pmid: 21816276
9. T. Porstmann *et al*., SREBP activity is regulated by mTORC1 and contributes to Akt-dependent cell growth. *Cell Metab.* **8**, 224–236 (2008). doi: 10.1016/j.cmet.2008.07.007; pmid: 18762023
10. H. L. Kenerson, M. M. Yeh, R. S. Yeung, Tuberous sclerosis complex-1 deficiency attenuates diet-induced hepatic lipid accumulation. *PLOS ONE* **6**, e18075 (2011). doi: 10.1371/journal.pone.0018075; pmid: 21479224
11. J. L. Yecies *et al*., Akt stimulates hepatic SREBP1c and lipogenesis through parallel mTORC1-dependent and independent pathways. *Cell Metab.* **14**, 21–32 (2011). doi: 10.1016/j.cmet.2011.06.002; pmid: 21723501
12. W. J. Quinn3rd *et al*., mTORC1 stimulates phosphatidylcholine synthesis to promote triglyceride secretion. *J. Clin. Invest.* **127**, 4207–4215 (2017). doi: 10.1172/JCI96036; pmid: 29035283
13. A. Umemura *et al*., Liver damage, inflammation, and enhanced tumorigenesis after persistent mTORC1 inhibition. *Cell Metab.* **20**, 133–144 (2014). doi: 10.1016/j.cmet.2014.05.001; pmid: 24910242
14. S. Wada *et al*., The tumor suppressor FLCN mediates an alternate mTOR pathway to regulate browning of adipose tissue. *Genes Dev.* **30**, 2551–2564 (2016). doi: 10.1101/gad.287953.116; pmid: 27913603
15. J. Li *et al*., Myeloid folliculin balances mTOR activation to maintain innate immunity homeostasis. *JCI Insight* **5**, 126939 (2019). doi: 10.1172/jci.insight.126939; pmid: 30843872
16. G. Napolitano *et al*., A substrate-specific mTORC1 pathway underlies Birt–Hogg–Dubé syndrome. *Nature* **585**, 597–602 (2020). doi: 10.1038/s41586-020-2444-0; pmid: 32612235
17. L. S. Schmidt, Birt–Hogg–Dubé syndrome: from gene discovery to molecularly targeted therapies. *Fam. Cancer* **12**, 357–364 (2013). doi: 10.1007/s10689-012-9574-y; pmid: 23108783
18. Z.-Y. Tsun *et al*., The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. *Mol. Cell* **52**, 495–505 (2013). doi: 10.1016/j.molcel.2013.09.016; pmid: 24095279
19. J. Betschinger *et al*., Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153**, 335–347 (2013). doi: 10.1016/j.cell.2013.03.012; pmid: 23582324
20. J. C. Kennedy *et al*., Loss of FLCN inhibits canonical WNT signaling via TFE3. *Hum. Mol. Genet.* **28**, 3270–3281 (2019). doi: 10.1093/hmg/ddz158; pmid: 31272105
21. J. C. Kennedy, D. Khabibullin, Y. Boku, W. Shi, E. P. Henske, New developments in the pathogenesis of pulmonary cysts in Birt–Hogg–Dubé Syndrome. *Semin. Respir. Crit. Care Med.* **41**, 247–255 (2020). doi: 10.1055/s-0040-1708500; pmid: 32279295
22. S.-B. Hong *et al*., Inactivation of the *FLCN* tumor suppressor gene induces TFE3 transcriptional activity by increasing its nuclear localization. *PLOS ONE* **5**, e15793 (2010). doi: 10.1371/journal.pone.0015793; pmid: 21209915
23. M. Yan *et al*., Chronic AMPK activation via loss of FLCN induces functional beige adipose tissue through PGC-1α/ERRα. *Genes Dev.* **30**, 1034–1046 (2016). doi: 10.1101/gad.281410.116; pmid: 27151976
24. M. Baba *et al*., Folliculin encoded by the *BHD* gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15552–15557 (2006). doi: 10.1073/pnas.0603781103; pmid: 17028174
25. D. Khabibullin *et al*., Folliculin regulates cell–cell adhesion, AMPK, and mTORC1 in a cell-type-specific manner in lung-derived cells. *Physiol. Rep.* **2**, e12107 (2014). doi: 10.14814/phy2.12107; pmid: 25121506
26. M. Yan *et al*., The tumor suppressor folliculin regulates AMPK-dependent metabolic transformation. *J. Clin. Invest.* **124**, 2640–2650 (2014). doi: 10.1172/JCI71749; pmid: 24762438
27. A. R. Tee, B. D. Manning, P. P. Roux, L. C. Cantley, J. Blenis, Tuberous sclerosis complex gene products, Tuberin and Hamartin, control mTOR signaling by acting as a GTPase-activating protein complex toward Rheb. *Curr. Biol.* **13**, 1259–1268 (2003). doi: 10.1016/S0960-9822(03)00506-2; pmid: 12906785
28. J. R. Clapper *et al*., Diet-induced mouse model of fatty liver disease and nonalcoholic steatohepatitis reflecting clinical disease progression and methods of assessment. *Am. J. Physiol. Gastrointest. Liver Physiol.* **305**, G483–G495 (2013). doi: 10.1152/ajpgi.00079.2013; pmid: 23886860
29. M. L. Boland *et al*., Towards a standard diet-induced and biopsy-confirmed mouse model of non-alcoholic steatohepatitis: Impact of dietary fat source. *World J. Gastroenterol.* **25**, 4904–4920 (2019). doi: 10.3748/wjg.v25.i33.4904; pmid: 31543682
30. X. Wang *et al*., Hepatocyte TAZ/WWTR1 promotes inflammation and fibrosis in nonalcoholic steatohepatitis. *Cell Metab.* **24**, 848–862 (2016). doi: 10.1016/j.cmet.2016.09.016; pmid: 28068223
31. R. L. Jacobs *et al*., Impaired *de novo* choline synthesis explains why phosphatidylethanolamine *N*-methyltransferase-deficient mice are protected from diet-induced obesity. *J. Biol. Chem.* **285**, 22403–22413 (2010). doi: 10.1074/jbc.M110.108514; pmid: 20452975
32. J. A. Martina *et al*., The nutrient-responsive transcription factor TFE3 promotes autophagy, lysosomal biogenesis, and clearance of cellular debris. *Sci. Signal.* **7**, ra9 (2014). doi: 10.1126/scisignal.2004754; pmid: 24448649
33. J. Xiong *et al*., *TFE3* alleviates hepatic steatosis through autophagy-induced lipophagy and *PGC1α*-mediated fatty acid β-oxidation. *Int. J. Mol. Sci.* **17**, 387 (2016). doi: 10.3390/ijms17030387; pmid: 26999124

34. J. C. Yoon *et al*., Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* **413**, 131–138 (2001). doi: 10.1038/35093050; pmid: 11557972
35. Y. Wang, J. Viscarra, S.-J. Kim, H. S. Sul, Transcriptional regulation of hepatic lipogenesis. *Nat. Rev. Mol. Cell Biol.* **16**, 678–689 (2015). doi: 10.1038/nrm4074; pmid: 26490400
36. M. Gao, D. Liu, The liver X receptor agonist T0901317 protects mice from high fat diet-induced obesity and insulin resistance. *AAPS J.* **15**, 258–266 (2013). doi: 10.1208/s12248-012-9429-3; pmid: 23180161
37. J. D. Horton, J. L. Goldstein, M. S. Brown, SREBPs: Activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J. Clin. Invest.* **109**, 1125–1131 (2002). doi: 10.1172/JCI0215593; pmid: 11994399
38. D. Yabe, M. S. Brown, J. L. Goldstein, Insig-2, a second endoplasmic reticulum protein that binds SCAP and blocks export of sterol regulatory element-binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12753–12758 (2002). doi: 10.1073/pnas.162488899; pmid: 12242332
39. G. Le Martelot *et al*., REV-ERBα participates in circadian SREBP signaling and bile acid homeostasis. *PLOS Biol.* **7**, e1000181 (2009). doi: 10.1371/journal.pbio.1000181; pmid: 19721697
40. D. Guan *et al*., Diet-induced circadian enhancer remodeling synchronizes opposing hepatic lipid metabolic processes. *Cell* **174**, 831–842.e12 (2018). doi: 10.1016/j.cell.2018.06.031; pmid: 30057115
41. C. R. Yellaturu, X. Deng, E. A. Park, R. Raghow, M. B. Elam, Insulin enhances the biogenesis of nuclear sterol regulatory element-binding protein (SREBP)-1c by posttranscriptional down-regulation of Insig-2A and its dissociation from SREBP cleavage-activating protein (SCAP)·SREBP-1c complex. *J. Biol. Chem.* **284**, 31726–31734 (2009). doi: 10.1074/jbc.M109.050914; pmid: 19759400
42. R. Papazyan *et al*., Physiological suppression of lipotoxic liver damage by complementary actions of HDAC3 and SCAP/SREBP. *Cell Metab.* **24**, 863–874 (2016). doi: 10.1016/j.cmet.2016.10.012; pmid: 27866836
43. C. Settembre *et al*., TFEB links autophagy to lysosomal biogenesis. *Science* **332**, 1429–1433 (2011). doi: 10.1126/science.1204592; pmid: 21617040
44. J. D. Horton, Y. Bashmakov, I. Shimomura, H. Shimano, Regulation of sterol regulatory element binding proteins in livers of fasted and refed mice. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5987–5992 (1998). doi: 10.1073/pnas.95.11.5987; pmid: 9600904
45. E. Steingrímsson *et al*., Mitf and Tfe3, two members of the Mitf-Tfe family of bHLH-Zip transcription factors, have important but functionally redundant roles in osteoclast development. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4477–4482 (2002). doi: 10.1073/pnas.072071099; pmid: 11930005
46. P. Angulo *et al*., Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* **149**, 389–397.e10 (2015). doi: 10.1053/j.gastro.2015.04.043; pmid: 25935633
47. M. Matsumoto *et al*., An improved mouse model that rapidly develops fibrosis in non-alcoholic steatohepatitis. *Int. J. Exp. Pathol.* **94**, 93–103 (2013). doi: 10.1111/iep.12008; pmid: 23305254
48. G. Wei *et al*., Comparison of murine steatohepatitis models identifies a dietary intervention with robust fibrosis, ductular reaction, and rapid progression to cirrhosis and cancer. *Am. J. Physiol. Gastrointest. Liver Physiol.* **318**, G174–G188 (2020). doi: 10.1152/ajpgi.00041.2019; pmid: 31630534
49. K. A. Lytle, D. B. Jump, Is Western diet-induced nonalcoholic steatohepatitis in *Ldlr*$^{-/-}$ mice reversible? *PLOS ONE* **11**, e0146942 (2016). doi: 10.1371/journal.pone.0146942; pmid: 26761430
50. C. D. Vocke *et al*., High frequency of somatic frameshift BHD gene mutations in Birt–Hogg–Dubé-associated renal tumors. *J. Natl. Cancer Inst.* **97**, 931–935 (2005). doi: 10.1093/jnci/dji154; pmid: 15956655
51. A. A. Barba, S. Bochicchio, A. Dalmoro, G. Lamberti, Lipid delivery systems for nucleic-acid-based-drugs: From production to clinical applications. *Pharmaceutics* **11**, 360 (2019). doi: 10.3390/pharmaceutics11080360; pmid: 31344836
52. A. J. Debacker, J. Voutila, M. Catley, D. Blakey, N. Habib, Delivery of oligonucleotides to the liver with GalNAc: From research to registered therapeutic drug. *Mol. Ther.* **28**, 1759–1771 (2020). doi: 10.1016/j.ymthe.2020.06.015; pmid: 32592692
53. J. Rüger, S. Ioannou, D. Castanotto, C. A. Stein, Oligonucleotides to the (gene) rescue: FDA approvals 2017–2019. *Trends Pharmacol. Sci.* **41**, 27–41 (2020). doi: 10.1016/j.tips.2019.10.009; pmid: 31836192
54. M. Baba *et al*., Kidney-targeted Birt-Hogg-Dubé gene inactivation in a mouse model: Erk1/2 and Akt-mTOR activation, cell hyperproliferation, and polycystic kidneys. *J. Natl. Cancer Inst.* **100**, 140–154 (2008). doi: 10.1093/jnci/djm288; pmid: 18182616
55. V. Hudon *et al*., Renal tumour suppressor function of the Birt–Hogg–Dubé syndrome gene product folliculin. *J. Med. Genet.* **47**, 182–189 (2010). doi: 10.1136/jmg.2009.072009; pmid: 19843504
56. A. Hagiwara *et al*., Hepatic mTORC2 activates glycolysis and lipogenesis through Akt, glucokinase, and SREBP1c. *Cell Metab.* **15**, 725–738 (2012). doi: 10.1016/j.cmet.2012.03.015; pmid: 22521878
57. M. Yuan, E. Pino, L. Wu, M. Kacergis, A. A. Soukas, Identification of Akt-independent regulation of hepatic lipogenesis by mammalian target of rapamycin (mTOR) complex 2. *J. Biol. Chem.* **287**, 29579–29588 (2012). doi: 10.1074/jbc.M112.386854; pmid: 22773877
58. N. Bhat *et al*., Dyrk1b promotes hepatic lipogenesis by bypassing canonical insulin signaling and directly activating mTORC2 in mice. *J. Clin. Invest.* **132**, 153724 (2022). pmid: 34855620
59. K. L. Donnelly *et al*., Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *J. Clin. Invest.* **115**, 1343–1351 (2005). doi: 10.1172/JCI23621; pmid: 15864352
60. J. E. Lambert, M. A. Ramos-Roman, J. D. Browning, E. J. Parks, Increased de novo lipogenesis is a distinct characteristic of individuals with nonalcoholic fatty liver disease. *Gastroenterology* **146**, 726–735 (2014). doi: 10.1053/j.gastro.2013.11.049; pmid: 24316260
61. C.-W. Kim *et al*., Acetyl CoA carboxylase inhibition reduces hepatic steatosis but elevates plasma triglycerides in mice and humans: A bedside to bench investigation. *Cell Metab.* **26**, 394–406.e6 (2017). doi: 10.1016/j.cmet.2017.07.009; pmid: 28768177
62. S. L. Friedman, B. A. Neuschwander-Tetri, M. Rinella, A. J. Sanyal, Mechanisms of NAFLD development and therapeutic strategies. *Nat. Med.* **24**, 908–922 (2018). doi: 10.1038/s41591-018-0104-9; pmid: 29967350
63. Y. Hu *et al*., Fructose and glucose can regulate mammalian target of rapamycin complex 1 and lipogenic gene expression via distinct pathways. *J. Biol. Chem.* **293**, 2006–2014 (2018). doi: 10.1074/jbc.M117.782557; pmid: 29222328

ACKNOWLEDGMENTS

We thank T. Harris for the Lipin1 antibodies; F. Foufelle for the INSIG2 antibody; L. Schmidt for the *Flcn*$^{lox/lox}$ mouse; D. Fisher for the *Tfe3*-null mouse; M. Giacca and L. Zentilin for the AAV-ApoE/AAT plasmid backbone; D. Cromley for Axcel analysis; L. Cheng and the University of Pennsylvania Cardiovascular Institute Histology Core for processing and staining of histology samples; the University of Pennsylvania Diabetes Research Center (DRC) for the use of the viral vector and the metabolomics cores (P30-DK19525); J. Li, I. Soaita, M. Blair, J. Axsom, M. Noji, and C. Bowman for technical assistance; Y. Kim, H. C. B. Nguyen, and M. Adlanmerini for expert advice on ChIP-seq; G. Liang for expert advice on nuclear SREBP-1 immunoblotting; M. Lazar for the AAV-nSREBP-1c plasmid; D. Salisbury for advice on LXR primers; P. Tontonoz and S. Lee for expert advice on T0901317 experiments; Penn Vector Core for production of AAVs; the Penn Rodent Metabolic Phenotyping Core for CLAMS studies; and the Penn Vet Comparative Pathology Core for blind scoring of mouse liver slides. **Funding:** This work was supported by an F30 NRSA fellowship from the NIDDK (F30 DK120096) and a Blavatnik Family Foundation Fellowship Award to B.S.G.; the DRC Regional Metabolomics Core (P30 DK19525); and NIH support for Z.A. (R01 DK107667). **Author contributions:** B.S.G. and Z.A. designed this study and wrote the manuscript. B.S.G. performed and/or contributed to all experiments. S.W. performed long-term AMLN diet experiments. C.T. contributed to AAV-HA-nSREBP-1c experiments. K.L. contributed to ChIP-seq experiments and helped with manuscript revisions. C.J. and S.Ju. performed mass spectrometry experiments and contributed to data analysis. Y.Y. and J.H.R. contributed to bioinformatics analysis. J.B. contributed to ChIP-seq experiments. S.Je. and L.L. contributed to histological experiments. K.U. and P.M.T. provided *Raptor* KO samples, and K.U. performed isolated hepatocyte assays. M.L. performed blinded histological evaluation. N.B.S. helped with general computational methods and assisted with analysis of CLAMS data. M.R.B. performed nonesterified fatty acid quantification for FPC diet. S.B.B. provided *Tsc1* KO samples and assisted with manuscript preparation. Z.A. oversaw the project. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Next-generation sequencing data are available through GEO accession number GSE160292.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abf8271
Materials and Methods
Figs. S1 to S12
Table S1
References (*64*–*89*)
MDAR Reproducibility Checklist

30 November 2020; accepted 23 February 2022
10.1126/science.abf8271

RESEARCH ARTICLE SUMMARY

CORONAVIRUS

KIR$^+$CD8$^+$ T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19

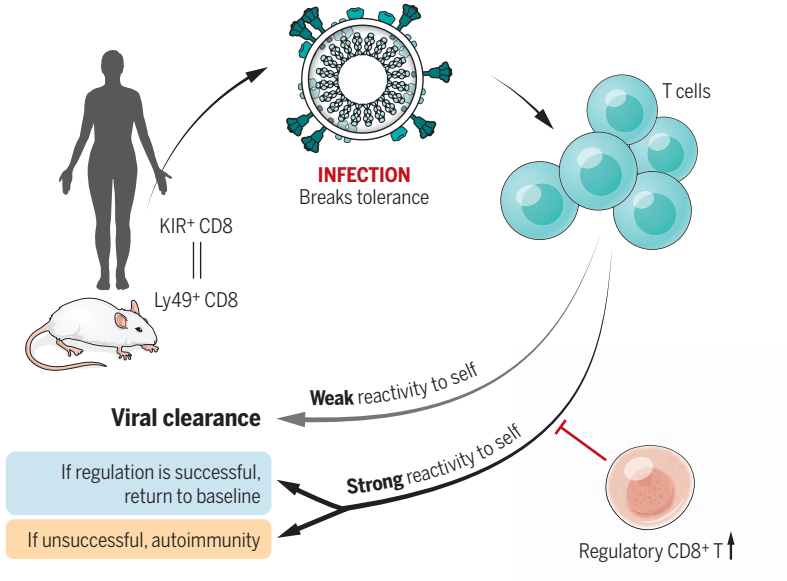
Jing Li, Maxim Zaslavsky, Yapeng Su, Jing Guo, Michael J. Sikora, Vincent van Unen, Asbjørn Christophersen, Shin-Heng Chiou, Liang Chen, Jiefu Li, Xuhuai Ji, Julie Wilhelmy, Alana M. McSween, Brad A. Palanski, Venkata Vamsee Aditya Mallajosyula, Nathan A. Bracey, Gopal Krishna R. Dhondalay, Kartik Bhamidipati, Joy Pai, Lucas B. Kipp, Jeffrey E. Dunn, Stephen L. Hauser, Jorge R. Oksenberg, Ansuman T. Satpathy, William H. Robinson, Cornelia L. Dekker, Lars M. Steinmetz, Chaitan Khosla, Paul J. Utz, Ludvig M. Sollid, Yueh-Hsiu Chien, James R. Heath, Nielsen Q. Fernandez-Becker, Kari C. Nadeau, Naresha Saligrama*, Mark M. Davis*

INTRODUCTION: Previous studies have identified Ly49$^+$CD8$^+$ T cells as a CD8$^+$ T cell subset with regulatory functions in mice. These cells can suppress myelin oligodendrocyte glycoprotein (MOG)–specific pathogenic CD4$^+$ T cells through their cytolytic activity and thereby ameliorate experimental autoimmune encephalomyelitis (EAE). However, whether a similar CD8$^+$ regulatory T cell subset exists in humans and whether its suppressive activity extends beyond autoimmune diseases to play a more general role in peripheral tolerance remains to be determined.

RATIONALE: Because killer cell immunoglobulin-like receptors (KIRs) are the evolutionary counterpart of the mouse Ly49 family in humans, we investigated whether KIR$^+$CD8$^+$ T cells are the phenotypic and functional equivalent of mouse Ly49$^+$CD8$^+$ T cells. We assessed the frequency of KIR$^+$CD8$^+$ T cells in human autoimmune and infectious diseases and analyzed their transcriptional profiles as well as T cell receptor (TCR) repertoires. Moreover, we developed an in vitro functional assay to test their regulatory functions on gliadin-specific disease-driving CD4$^+$ T cells from patients with celiac disease (CeD) and to study the mechanisms of suppression. To further elucidate their role in infectious diseases, we analyzed the effects of selective ablation of their murine counterpart on the antiviral responses during infection and tissue pathology after resolution of viral infection.

RESULTS: The frequency of KIR$^+$CD8$^+$ T cells was higher in the blood and inflamed tissues of patients with multiple autoimmune diseases compared with those of healthy controls. Furthermore, the increase of duodenal KIR$^+$CD8$^+$ T cells positively correlated with disease activity in CeD. In vitro, KIR$^+$CD8$^+$ T cells were able to specifically eliminate gliadin-specific pathogenic CD4$^+$ T cells from the leukocytes of CeD patients through their cytotoxicity in a class I major histocompatibility complex (MHC)–dependent manner. Moreover, our RNA sequencing (RNA-seq) analysis revealed many similarities between human KIR$^+$CD8$^+$ T cells and mouse Ly49$^+$CD8$^+$ T cells, which suggests that KIR$^+$CD8$^+$ T cells are the functional and phenotypic equivalents of mouse Ly49$^+$CD8$^+$ T cells in humans. The expression of inhibitory KIR receptors appeared to suppress the activation and cytotoxic functions of KIR$^+$CD8$^+$ T cells. Moreover, elevated levels of KIR$^+$CD8$^+$ T cells, but not CD4$^+$ regulatory T cells (T_{regs}), were also observed in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)– or influenza-infected patients and correlated with vasculitis in COVID-19 patients. Notably, expanded KIR$^+$CD8$^+$ T cells from healthy subjects and different diseases displayed shared phenotypes and similar TCR sequences as revealed by single-cell RNA-seq and TCR sequencing (TCR-seq). Finally, the selective depletion of Ly49$^+$CD8$^+$ T cells in virus-infected mice did not interfere with the antiviral responses but resulted in exacerbated autoimmunity after infection, demonstrating the critical role that this subset of CD8$^+$ T_{regs} plays in controlling harmful self-reactivity in infectious diseases.

CONCLUSION: We identify KIR$^+$CD8$^+$ T cells as an important regulatory T cell subset in humans. They are induced as part of the response during an autoimmune reaction or infection and may act as a negative feedback mechanism to specifically suppress the self-reactive or otherwise pathogenic cells without affecting the immune responses against pathogens. This subset of CD8$^+$ T_{regs} appears to play an important role in maintaining peripheral tolerance, which is distinct from and likely complementary to that of CD4$^+$ T_{regs}. Our findings also provide insights into understanding the relationship between autoimmunity and infectious diseases and into the development of potential therapeutic approaches targeting KIR$^+$CD8$^+$ T cells to suppress undesirable self-reactivity in autoimmune disorders and infectious diseases. ■



The proposed role of CD8$^+$ regulatory T cells in peripheral tolerance. KIR$^+$CD8$^+$ T cells are the equivalent of mouse Ly49$^+$CD8$^+$ T cells in humans, with similar regulatory functions. During an infection, this kind of CD8$^+$ regulatory T cell is induced to suppress those CD4$^+$ T cells with a strong reactivity to self, which may cause autoimmunity, without interfering with the immune responses against pathogens.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: nareshas@wustl.edu (N.S.); mmdavis@stanford.edu (M.M.D.)
This is an open-access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Cite this article as J. Li *et al.*, *Science* **376**, eabi9591 (2022). DOI: 10.1126/science.abi9591

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abi9591

RESEARCH ARTICLE

CORONAVIRUS

KIR$^+$CD8$^+$ T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19

Jing Li[1], Maxim Zaslavsky[2], Yapeng Su[3], Jing Guo[4], Michael J. Sikora[5], Vincent van Unen[1], Asbjørn Christophersen[6,7,8], Shin-Heng Chiou[1]†, Liang Chen[1], Jiefu Li[9], Xuhuai Ji[10], Julie Wilhelmy[1], Alana M. McSween[1]‡, Brad A. Palanski[11]§, Venkata Vamsee Aditya Mallajosyula[1], Nathan A. Bracey[1], Gopal Krishna R. Dhondalay[12], Kartik Bhamidipati[13], Joy Pai[13], Lucas B. Kipp[14], Jeffrey E. Dunn[14], Stephen L. Hauser[15], Jorge R. Oksenberg[15], Ansuman T. Satpathy[16], William H. Robinson[17,18], Cornelia L. Dekker[19], Lars M. Steinmetz[5,20,21], Chaitan Khosla[11,22,23], Paul J. Utz[1,18], Ludvig M. Sollid[6,7,8,24], Yueh-Hsiu Chien[1,4], James R. Heath[3,25], Nielsen Q. Fernandez-Becker[26], Kari C. Nadeau[1,12], Naresha Saligrama[1]*¶, Mark M. Davis[1,9,27]*

In this work, we find that CD8$^+$ T cells expressing inhibitory killer cell immunoglobulin-like receptors (KIRs) are the human equivalent of Ly49$^+$CD8$^+$ regulatory T cells in mice and are increased in the blood and inflamed tissues of patients with a variety of autoimmune diseases. Moreover, these CD8$^+$ T cells efficiently eliminated pathogenic gliadin-specific CD4$^+$ T cells from the leukocytes of celiac disease patients in vitro. We also find elevated levels of KIR$^+$CD8$^+$ T cells, but not CD4$^+$ regulatory T cells, in COVID-19 patients, correlating with disease severity and vasculitis. Selective ablation of Ly49$^+$CD8$^+$ T cells in virus-infected mice led to autoimmunity after infection. Our results indicate that in both species, these regulatory CD8$^+$ T cells act specifically to suppress pathogenic T cells in autoimmune and infectious diseases.

Although most CD8$^+$ T cells are geared toward the control of pathogen-infected or cancerous cells, there has been long-standing evidence in mice that a small subset can also suppress autoimmune responses (*1*). This regulatory function of CD8$^+$ T cells was first implicated by the depletion of CD8$^+$ T cells in experimental autoimmune encephalomyelitis (EAE), a mouse model of human multiple sclerosis (MS) (*2*, *3*). In particular, disruption of Qa-1-CD8 co-receptor binding in B6.Qa-1-D227K mice leads to spontaneous autoimmune diseases (*4*). The Ly49 family of inhibitory C-type lectin-like receptors, which are ubiquitous on natural killer (NK) cells, were identified as specific surface markers for this regulatory CD8$^+$ T cell subset (*5*). Moreover, the transcription factor Helios is an essential control element for their differentiation and function in mice (*6*). Recently, our research group found that clonally expanded CD8$^+$ T cells in EAE recognized peptides bound to H2-D^b and that these peptides stimulate Ly49$^+$CD8$^+$ regulatory T cells and suppress disease (*7*). This extended the original observations beyond Qa-1 to encompass classical class I major histocompatibility complex (MHC) interactions, suggesting a general mechanism of peripheral tolerance. In this work, we identify CD8$^+$ T cells expressing inhibitory killer cell immunoglobulin-like receptors (KIRs)—the functional counterpart of the mouse Ly49 family in humans (*8*)—as a regulatory CD8$^+$ T cell subset in humans that suppresses pathogenic CD4$^+$ T cells in celiac disease (CeD) and likely other autoimmune disorders and infectious diseases as well.

Increased KIR$^+$CD8$^+$ T cells in human autoimmune diseases

Both mouse Ly49 and human KIR receptors are known to bind to class I MHC molecules (*8*). They typically have immunoreceptor tyrosine-based inhibitory motifs (ITIMs) in their cytoplasmic tails and are ubiquitously expressed on NK cells as well as on a small subset (1 to 5%) of CD8$^+$ T cells (*5*). Therefore, we analyzed CD8$^+$ T cells expressing inhibitory KIRs (which we refer to as KIR$^+$CD8$^+$ T cells) (*9*, *10*) in the peripheral blood of patients with autoimmune diseases and age- and gender-matched healthy controls (HCs). KIR3DL1 and KIR2DL3 were the two major KIR subtypes expressed by human CD8$^+$ T cells (fig. S1). The frequency of KIR$^+$CD8$^+$ T cells was significantly increased in the blood of patients with MS, systemic lupus erythematosus (SLE), or CeD compared with the blood of HCs (Fig. 1A).

Next, we investigated whether KIR$^+$CD8$^+$ T cells were also present in the inflamed tissues of patients with these diseases. CeD patients with active disease had higher levels of KIR$^+$CD8$^+$ T cells in the gut compared with those in remission (on a gluten-free diet) as well as the non-CeD controls (Fig. 1B), which indicates a strong correlation of KIR$^+$CD8$^+$ T cell levels with disease severity. Additionally, the number of KIR$^+$CD8$^+$ T cells was markedly increased in the kidneys of patients with SLE compared with healthy kidneys (Fig. 1C) and in the synovial tissues of rheumatoid arthritis (RA) patients compared with those with osteoarthritis (OA), which is not thought to be an autoimmune disease. By contrast, the frequencies of synovial FOXP3$^+$CD4$^+$ regulatory T cells (T_{regs}) were similar between RA and OA patients (Fig. 1D).

KIR$^+$CD8$^+$ T cells are the functional and phenotypic equivalent of mouse Ly49$^+$CD8$^+$ T cells

We next investigated whether KIR$^+$CD8$^+$ T cells are the functional counterpart of mouse Ly49$^+$ regulatory CD8$^+$ T cells. Previously we had found that Ly49$^+$CD8$^+$ T cells suppress myelin oligodendrocyte glycoprotein (MOG)–specific pathogenic CD4$^+$ T cells in a perforin-dependent manner (*4*, *7*). Deamidated gliadin derived from dietary gluten is the antigen for CD4$^+$ T cells that drives autoimmune enteropathy in human CeD (*11*, *12*). Therefore, we explored whether KIR$^+$CD8$^+$ T cells can suppress

[1]Institute of Immunity, Transplantation and Infection, Stanford University School of Medicine, Stanford, CA, USA. [2]Program in Computer Science, Stanford University, Stanford, CA, USA. [3]Institute for Systems Biology, Seattle, WA, USA. [4]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA, USA. [5]Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. [6]K.G. Jebsen Coeliac Disease Research Centre, University of Oslo, Oslo, Norway. [7]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [8]Department of Immunology, University of Oslo, Oslo, Norway. [9]The Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA. [10]Human Immune Monitoring Center, Stanford University School of Medicine, Stanford, CA, USA. [11]Department of Chemistry, Stanford University, Stanford, CA, USA. [12]Sean N. Parker Center for Allergy and Asthma Research, Department of Medicine, Stanford University, Stanford, CA, USA. [13]Program in Immunology, Stanford University School of Medicine, Stanford, CA, USA. [14]Division of Neuroimmunology, Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA. [15]Department of Neurology and UCSF Weill Institute for Neurosciences, University of California, San Francisco, CA, USA. [16]Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA. [17]VA Palo Alto Health Care System, Palo Alto, CA, USA. [18]Division of Immunology and Rheumatology, Department of Medicine, Stanford University, Stanford, CA, USA. [19]Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA. [20]Stanford Genome Technology Center, Stanford University, Palo Alto, CA, USA. [21]European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Heidelberg, Germany. [22]Department of Chemical Engineering, Stanford University, Stanford, CA, USA. [23]Stanford ChEM-H, Stanford University, Stanford, CA, USA. [24]Department of Immunology, Oslo University Hospital, Oslo, Norway. [25]Department of Bioengineering, University of Washington, Seattle, WA, USA. [26]Division of Gastroenterology and Hepatology, Department of Medicine, Stanford University, Stanford, CA, USA. [27]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA, USA.

***Corresponding author. Email: nareshas@wustl.edu (N.S.); mmdavis@stanford.edu (M.M.D.)**

†Present address: Rutgers Cancer Institute of New Jersey, Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA. ‡Present address: Immunology Graduate Program, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA. §Present address: Department of Medicine, Division of Genetics, Brigham and Women's Hospital, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. ¶Present address: Department of Neurology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA.

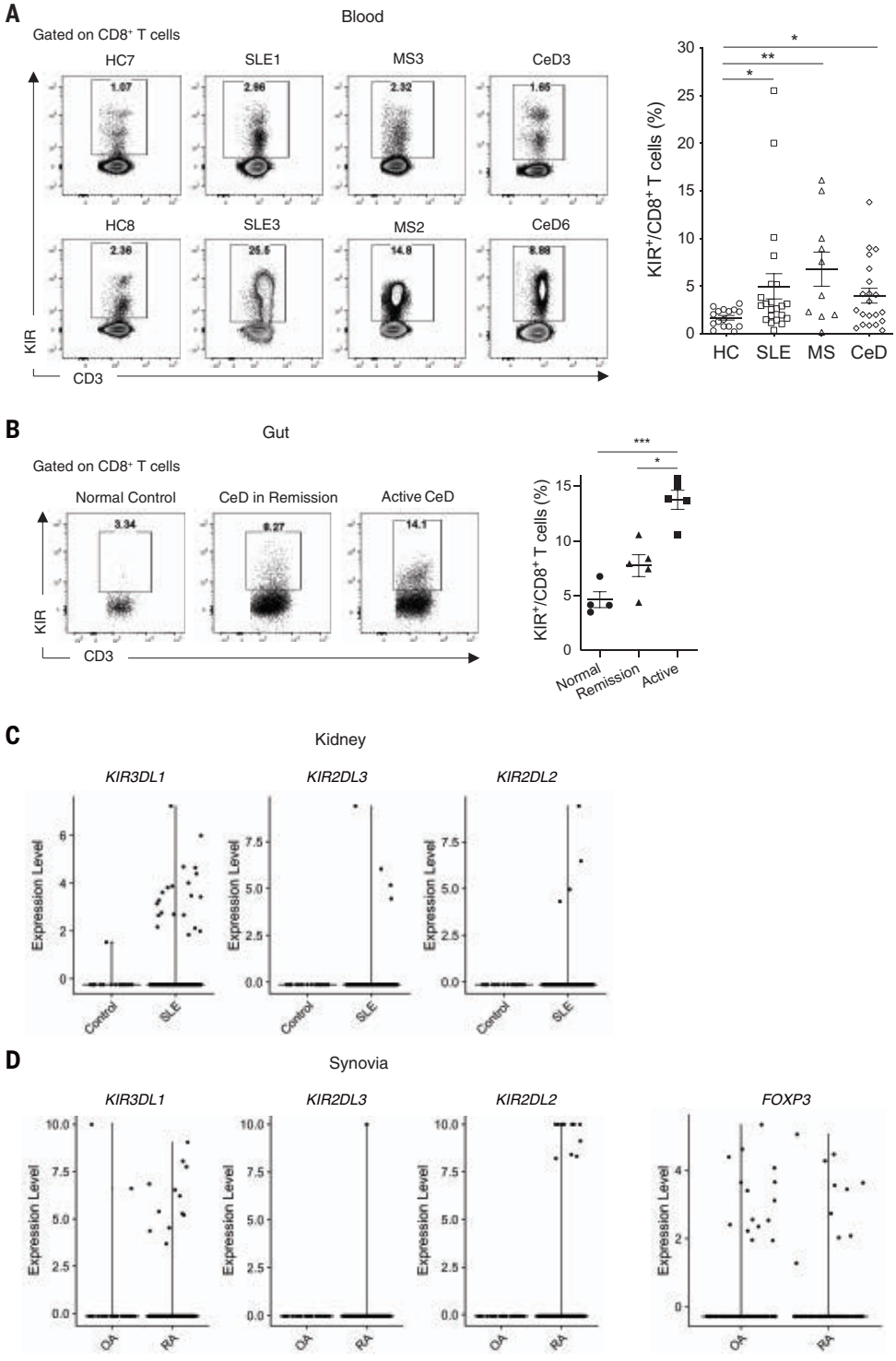


Fig. 1. Increased KIR$^+$CD8$^+$ T cells in patients with autoimmune diseases. (**A**) Representative contour plots and a summary histogram showing the frequency of KIR$^+$CD8$^+$ T cells in the peripheral blood of HCs (N = 16) and patients with SLE (N = 22), MS (N = 10), or CeD (N = 21) analyzed by flow cytometry. KIR$^+$ cells were detected by PE-conjugated antibodies against KIR2DL1 (clone no. 143211), KIR2DL2/L3 (Dx27), KIR2DL5 (UP-R1), KIR3DL1 (Dx9), and KIR3DL2 (clone no. 539304). *P < 0.05; **P < 0.01; Kruskal-Wallis test followed by multiple comparisons test. (**B**) Representative contour plots and a summary histogram showing the frequency of KIR$^+$CD8$^+$ T cells in the duodenum of normal controls (N = 4), CeD in remission (N = 5), and active CeD (N = 5) analyzed by flow cytometry. *P < 0.05; ***P < 0.001; Kruskal-Wallis test followed by multiple comparisons test. (**C**) Expression of *KIR* transcripts (*KIR3DL1*, *KIR2DL3*, and *KIR2DL2*) in CD8$^+$ T cells from healthy kidneys (control, N = 6) versus SLE nephritis kidneys (N = 20). (**D**) Expression of *KIR* transcripts (*KIR3DL1*, *KIR2DL3*, and *KIR2DL2*) in synovial CD8$^+$ T cells and expression of *FOXP3* in synovial CD4$^+$ T cells from RA (N = 18) and OA (N = 3).

gliadin-specific CD4$^+$ T cells from CeD patients. CD8$^-$ peripheral blood mononuclear cells (PBMCs) from human leukocyte antigen (HLA)–DQ2.5$^+$ CeD patients were cultured with preactivated KIR$^+$ and KIR$^-$CD8$^+$ T cells supplemented with deamidated gluten (fig. S2A). In the absence of KIR$^+$CD8$^+$ T cells, deamidated gluten profoundly stimulated the expansion of gliadin-specific CD4$^+$ T cells. Notably, stimulated KIR$^+$CD8$^+$ T cells, but not KIR$^-$CD8$^+$ T cells or KIR$^+$ NK cells, significantly reduced the number of gliadin-specific CD4$^+$ T cells (Fig. 2A and fig. S2C) without affecting the number of total CD4$^+$ T cells (fig. S2B). KIR$^+$CD8$^+$ T cells appeared to target only the pathogenic CD4$^+$ T cells because they had no discernible effect on hemagglutinin (HA)–specific CD4$^+$ T cells induced by influenza A HA peptides (fig. S2D) or on the proliferation of CD4$^+$ T cells after anti-CD3 stimulation (fig. S2E). The suppression by KIR$^+$CD8$^+$ T cells was contact dependent because their inhibitory effects on gliadin-specific CD4$^+$ T cells were abrogated when they were separated from the CD8$^-$ PBMCs by a membrane insert (Fig. 2C). We also found increased annexin V binding on gliadin-specific CD4$^+$ T cells in the presence of KIR$^+$CD8$^+$ T cells

(Fig. 2B), indicating that these cells can induce apoptosis of the pathogenic $CD4^+$ T cells. In the presence of high-dose interleukin-2 (IL-2), KIR^+CD8^+ T cells were still able to reduce the number of gliadin-specific $CD4^+$ T cells (fig. S2F). Thus, KIR^+CD8^+ T cells suppress pathogenic $CD4^+$ T cells through direct killing instead of a competition for IL-2, consistent with the perforin dependance of $Ly49^+CD8^+$ T cells in mice (*4*, *7*).

Previous studies have suggested that the regulatory function of mouse $Ly49^+CD8^+$ T cells is mediated by recognition of both classical (*7*) and nonclassical class I MHCs (*4*, *5*, *13*) on their target cells. Preactivated KIR^+CD8^+ T cells showed more-potent suppression of gliadin-specific $CD4^+$ T cells than the untreated KIR^+CD8^+ T cells (Fig. 2D), which indicates that T cell receptor (TCR) activation is required to fully elicit their suppressive functions. Moreover, anti–HLA-ABC and anti–HLA-E blockade could partially reverse the suppression by KIR^+CD8^+ T cells (Fig. 2E). Thus, KIR^+CD8^+ T cells specifically eliminate gliadin-specific $CD4^+$ T cells from the leukocytes of CeD patients through the recognition of classical and/or nonclassical class I MHC molecules.

To further investigate whether KIR^+CD8^+ T cells are the phenotypic equivalent of mouse $Ly49^+$ T cells in humans, we performed RNA sequencing (RNA-seq) analysis on KIR^+ versus KIR^-CD8^+ T cells from patients with MS to compare with mouse $Ly49^+CD8^+$ T cells in EAE, a mouse model of human MS. There were 778 differentially expressed genes (DEGs) between KIR^+ and KIR^-CD8^+ T cells (table S1). Notably, KIR^+CD8^+ T cells showed a marked up-regulation of cytotoxic molecules, NK cell–associated genes, and cell-trafficking molecules, in addition to inhibitory KIR receptor genes (fig. S3A). Furthermore, KIR^+CD8^+ T cells expressed higher transcript levels for Helios (encoded by *IKZF2*), a transcription factor associated with regulatory functions of both $CD4^+$ and $CD8^+$ T cells (*6*). KIR^+CD8^+ T cells also down-regulated naïve and memory T cell–associated molecules and the costimulatory receptor *CD28* (fig. S3A), which is one of the key features for regulatory $CD8^+$ T cell populations in mice and humans (*14*). Gene ontology enrichment analysis of these DEGs showed enrichment for T cell activation, proliferation, migration, and differentiation (fig. S3B). Moreover, gene set enrichment analysis (GSEA) (*15*, *16*) revealed that approximately half of the top 200 genes up-regulated in $Ly49^+CD8^+$ T cells were also higher in KIR^+ $CD8^+$ T cells (fig. S3C). Previously, we found that $Ly49^+CD8^+$ T cells expressed 16 of the 60 genes conserved in $CD4^+$ T_{regs} (*7*). These same T_{reg} signature genes (*17*) were also enriched in KIR^+CD8^+ T cells (fig. S3D). Thus, human KIR^+CD8^+ T cells share many similarities with $Ly49^+CD8^+$ T cells from EAE mice.

RNA-seq analysis of KIR^+ versus KIR^-CD8^+ T cells from healthy subjects and patients with different autoimmune diseases (including MS, SLE, and CeD) identified a set of 963 DEGs. Many of them overlapped with the DEGs previously defined in MS (fig. S3E and table S2), but larger-fold changes were observed in patients with higher frequencies of KIR^+CD8^+ T cells (fig. S4A). Consistent with the transcriptional profiles, KIR^+CD8^+ T cells had higher protein expression of granzyme B, perforin, CX3CR1, KLRG1, CD244, TIGIT, T-bet, and Helios and lower levels of CCR7, CD27, and CD28 (fig. S4B). Similar to the circulating KIR^+CD8^+ T cells, both kidney and synovial KIR^+CD8^+ T cells up-regulated *KLRG1*, *CD244*, *TIGIT*, *CX3CR1*, *PRF1*, *GZMB*, and *IKZF2* and down-regulated *CD28* and *CCR7* (fig. S5). Thus, KIR^+CD8^+ T cells appear to be the functional and phenotypic equivalent of mouse $Ly49^+CD8^+$ T cells in humans, with many conserved features in both healthy subjects and those with autoimmune diseases.

Because the KIR^+CD8^+ T cells we focused on in this study express inhibitory KIR receptors, which contain intracellular ITIMs, we investigated how these KIR receptors contribute to the differentiation and functionality of KIR^+CD8^+ T cells. $CD8^+$ T cells with low KIR3DL1 or KIR2DL3 expression displayed higher surface CX3CR1 and intracellular granzyme B and perforin compared with those with high KIR expression (fig. S6, A to B). RNA-seq followed by gene ontology enrichment analysis confirmed that $CD8^+$ T cells with low KIR expression displayed enhanced cytotoxicity and T cell activation (fig. S6, C to D). Thus, the high expression of inhibitory KIR receptors suppresses the activation and effector functions of the KIR^+CD8^+ T cells. This may allow for the precise control of their activity to avoid bystander suppression.

Increased KIR^+CD8^+ T cells in SARS-CoV-2– and influenza-infected patients

Although it was previously accepted that most self-specific T cells were eliminated in the thymus, recent work has shown that many such cells survive and populate the periphery of both humans and mice (*18*, *19*). We have hypothesized that this occurs because the threat of infectious diseases (*20*) necessitates a complete T cell repertoire (*18*, *21*), such that even self-reactive T cells might be needed in the response to a particular pathogen. Consistent with this are classic experiments showing that infectious diseases or treatments that mimic them can activate self-specific T cells (*22*). This led us to hypothesize that KIR^+CD8^+ T cells might be elevated during an infection to control autoreactive T cells. We first analyzed 53 patients with COVID-19. We found that KIR^+CD8^+ T cells were substantially elevated in many of these patients and higher levels correlated with more severe disease (Fig. 3A). Moreover, the highest frequencies of KIR^+CD8^+ T cells were found in patients with vasculitis or embolism and, to a lesser extent, in those with acute respiratory distress syndrome (ARDS) (Fig. 3B and fig. S7, C to D), which are common complications of this disease, likely caused by excessive inflammation. Thus, increased levels of KIR^+CD8^+ T cells may be prognostic for autoimmune-related pathologies during a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. By contrast, we did not observe any difference in the levels of $CD25^{hi}CD127^{lo}CD4^+$ T_{regs} or KIR^+ NK cells in COVID-19 patients compared with HCs or in COVID-19 patients with different disease severities or complications (fig. S7, A to D, and Fig. 3B). Using publicly available single-cell RNA sequencing (scRNA-seq) data (*23*), we also found an increased frequency of KIR^+CD8^+ T cells in the bronchoalveolar lavage fluid of COVID-19 patients with moderate or severe disease compared with that from HCs (Fig. 3C). Additionally, an increased frequency of KIR^+CD8^+ T cells, but not $CD4^+$ T_{regs}, was observed in the peripheral blood of influenza-infected patients (Fig. 3D), which suggests that KIR^+CD8^+ T cells are generally induced as part of an infectious disease response.

Commonality and heterogeneity of KIR^+CD8^+ T cells

We next performed scRNA-seq analysis on peripheral blood $CD8^+$ T cells from HCs, MS patients, and COVID-19 patients (*24*) using the 10X Genomics platform (*25*). KIR^+CD8^+ T cells from different conditions formed a distinct cluster with high expression of effector genes as well as *KIR* transcripts (Fig. 4, A to B, and table S3). Compared with KIR^- effector $CD8^+$ T cells, KIR^+ effector $CD8^+$ T cells expressed higher levels of *IKZF2* and NK cell–associated genes (e.g., *TYROBP*, *KLRC2*, *KLRC3*, *NCR1*, and *NCR3*) while expressing lower levels of *IL7R*, *CD27*, and *KLRB1* (table S3). Thus, these findings reveal the commonality of KIR^+CD8^+ T cells across physiological and diseased statuses as well as their specialness relative to other $CD8^+$ T cells.

To better understand the similarity and heterogeneity of KIR^+CD8^+ T cells under different circumstances, we performed scRNA-seq on 4512 KIR^+CD8^+ T cells sorted from healthy subjects and patients with MS, SLE, or CeD using the Smart-seq2 protocol (*26*) and analyzed their TCR α and β sequences (*27*). Unsupervised clustering of these KIR^+CD8^+ T cells identified six clusters, with clusters 1 to 3 mostly containing expanded KIR^+CD8^+ T cells (≥2 cells expressing same TCR), and clusters 5 and 6 consisting of unexpanded cells

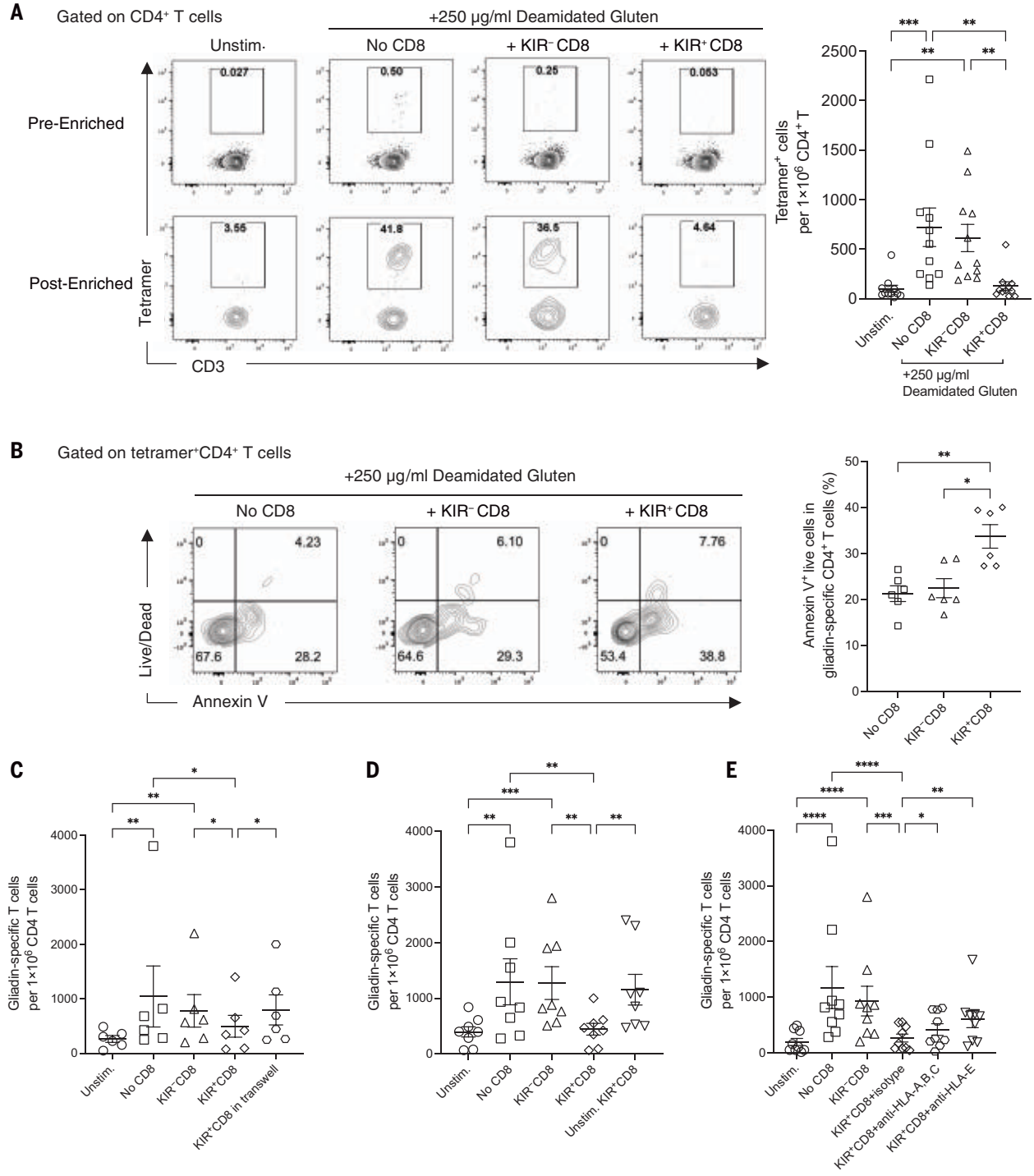


Fig. 2. Elimination of gliadin-specific CD4$^+$ T cells by KIR$^+$CD8$^+$ T cells. (**A**) Representative contour plots showing tetramer-bound CD4$^+$ T cells before and after enrichment and summary of the number of gliadin-specific CD4$^+$ T cells per 1 × 10^6 CD4$^+$ T cells on day 6. Experiments were repeated using PBMCs from 11 CeD patients. $^{**}P < 0.01$; $^{***}P < 0.001$; Friedman test followed by multiple comparisons test. Unstim., unstimulated. (**B**) Representative contour plots and a summary graph displaying annexin V binding of gliadin-specific CD4$^+$ T cells from the culture harvested on day 3 ($N = 6$). $^{*}P < 0.05$; $^{**}P < 0.01$; Friedman test followed by multiple comparisons test. (**C**) Frequency of gliadin-specific CD4$^+$ T cells from the cell cultures in the presence or absence of KIR$^-$ or KIR$^+$CD8$^+$ T cells or with KIR$^+$CD8$^+$ T cells separated by a 4-µm insert in a transwell plate ($N = 6$). $^{*}P < 0.05$; $^{**}P < 0.01$; Friedman test followed by multiple comparisons test. (**D**) Frequency of gliadin-specific CD4$^+$ T cells from the PBMC cultures in the presence or absence of KIR$^-$ or KIR$^+$CD8$^+$ T cells with or without preactivation ($N = 8$). $^{**}P < 0.01$; $^{***}P < 0.001$; Friedman test followed by multiple comparisons test. (**E**) Frequency of gliadin-specific CD4$^+$ T cells from the PBMC cultures in the presence or absence of preactivated KIR$^-$ or KIR$^+$CD8$^+$ T cells with isotype control, anti–HLA-ABC, or anti–HLA-E blockade ($N = 9$). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; Friedman test followed by multiple comparisons test.

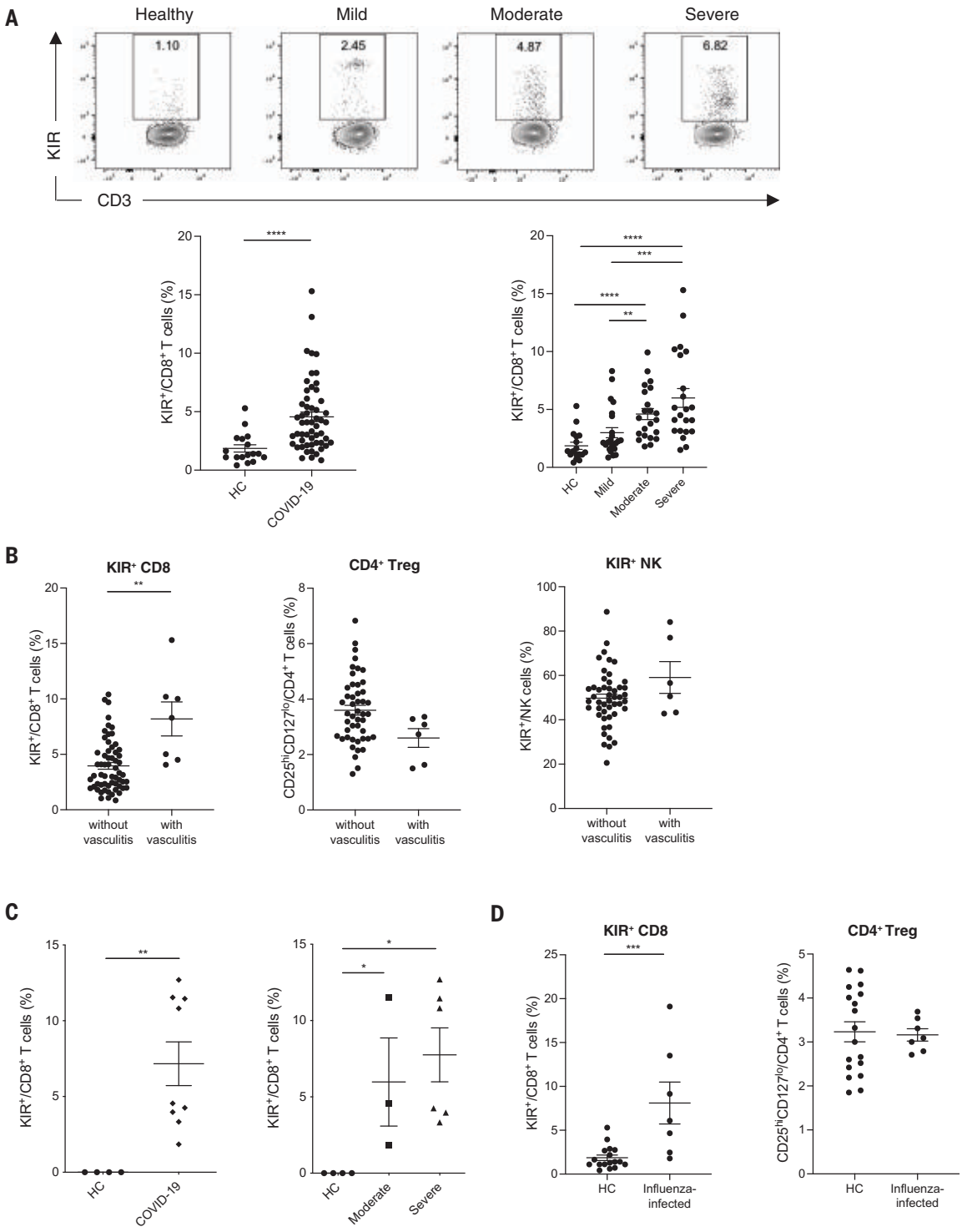


Fig. 3. Increased KIR$^+$CD8$^+$ T cells in infectious diseases. (**A**) Representative contour plots and summary scatter plots showing the percentage of KIR$^+$ cells in CD8$^+$ T cells from the blood of 17 HCs and 53 COVID-19 patients with varying disease severity (mild, *N* = 23; moderate, *N* = 17; severe, *N* = 13). ****P < 0.0001; Mann-Whitney test (left). **P < 0.01; ***P < 0.001; ****P < 0.0001; Kruskal-Wallis test followed by multiple comparisons test (right). (**B**) Frequency of KIR$^+$CD8$^+$ T cells, CD4$^+$ T_{regs} (CD25^{hi}CD127^{lo}), and KIR$^+$ NK cells in the blood of COVID-19 patients with or without vasculitis. **P < 0.01; Mann-Whitney test. (**C**) Frequency of CD8$^+$ T cells expressing *KIR* transcripts (*KIR3DL1*, *KIR3DL2*, *KIR2DL3*, or *KIR2DL1*) in the bronchoalveolar lavage fluid of HCs (*N* = 4) versus COVID-19 patients (*N* = 9) (left; **P < 0.01; Mann-Whitney test) and HCs versus COVID-19 patients with moderate (*N* = 3) or severe (*N* = 6) disease (right; *P < 0.05; Kruskal-Wallis test followed by multiple comparisons test). (**D**) Frequency of KIR$^+$CD8$^+$ T cells and CD4$^+$ T_{regs} (CD25^{hi}CD127^{low}) in the blood of HCs (*N* = 17) versus influenza-infected patients (*N* = 7). ***P < 0.001; Mann-Whitney test.

expressing unique TCRs (Fig. 4C and fig. S8, A to B). Expanded KIR$^+$ cells in clusters 1 to 3 had higher transcripts for cytotoxic molecules (e.g., *GZMH*, *GZMB*, and *PRF1*) and genes associated with effector T cells (e.g., *FCGR3A*, *FGFBP2*, and *CX3CR1*). Cluster 2, which was more restricted to expanded KIR$^+$ cells from MS patients, showed higher levels of type I interferon (IFN)–responding genes. Cluster 3—specific to expanded KIR$^+$ cells from a subset of HCs and SLE patients—showed higher expression of genes involved in glycolysis (Fig. 4, D to F, and table S4). Cells in cluster 4 were in a transitional state, with a loss of memory-associated features. Clusters 5 and 6 displayed memory and naïve signatures, respectively (Fig. 4F and table S4), and accounted for a small proportion of total KIR$^+$CD8$^+$ T cells (fig. S8A). T cell clones expressing identical TCRs could be found in different clusters, indicating possible lineage relationships. Additionally, clonally expanded KIR$^+$CD8$^+$ T cells in COVID-19 patients identified from the previous 10X Genomics scRNA-seq (*24*) displayed a higher expression of cytotoxic genes while down-regulating naïve- or memory-associated genes compared with unexpanded KIR$^+$CD8$^+$ T cells (fig. S8, C to D). In the assay with celiac PBMCs, CCR7$^-$ effector KIR$^+$CD8$^+$ T cells displayed stronger suppressive activity against gliadin-specific CD4$^+$ T cells than CCR7$^+$KIR$^+$CD8$^+$ T cells (fig. S2G), which was consistent with their up-regulated cytotoxic functions as revealed by the scRNA-seq. Thus, in parallel with clonal expansion, KIR$^+$CD8$^+$ T cells may lose their naïve or memory attributes, enter the differentiation program for effector T cells, and then suppress pathogenic CD4$^+$ T cells through cytotoxicity. There are common features shared by KIR$^+$CD8$^+$ T cells from healthy subjects and from subjects with different diseases, yet there is also heterogeneity associated with different diseases or treatments.

We next compared the diversity of KIR$^+$CD8$^+$ TCRs with KIR$^-$CD8$^+$ TCRs from the same individuals (N = 26) and found that TCRs of KIR$^+$CD8$^+$ T cells had significantly lower Shannon-Wiener indices and Chao estimates than KIR$^-$CD8$^+$ T cells (Fig. 5, A to B). Thus, the TCR repertoire of KIR$^+$CD8$^+$ T cells is less diverse, consistent with a previous study that KIR$^+$CD8$^+$ T cells display a more restricted TCR Vβ chain usage (*9*). To compare the antigen specificities of KIR$^+$CD8$^+$ T cells from different disease types, we also analyzed the TCR sequences using GLIPH2 (*28*), which is an algorithm to cluster TCRs that recognize the same antigen in most cases. TCRs of KIR$^+$CD8$^+$ T cells from healthy donors and different diseases could be grouped into the same GLIPH clusters, although with different extents of clonal expansion (Table 1), which indicates that they may recognize the same antigens that commonly exist under physiological and different pathological conditions. Thus, expanded KIR$^+$CD8$^+$ T cells have shared phenotypes and antigen specificity independent of disease type. Although the analysis above shows that the TCR repertoire of KIR$^+$CD8$^+$ T cells is less diverse than KIR$^-$CD8$^+$ T cells generally, it is still considerably diverse, using multiple classical HLAs and HLA-E and probably recognizing many antigenic peptides as well.

Regulatory CD8$^+$ T cells suppress autoimmunity developed after virus infection

We next studied the effect of selective ablation of Ly49$^+$CD8$^+$ T cells on virus-infected mice. Because Ly49F (encoded by *Klra6*) is expressed on 90% of Ly49$^+$CD8$^+$ T cells but only a small fraction of NK cells (*5*, *29*), we generated a *Klra6*^{cre} mouse line. In *Klra6*^{cre}R26R-EYFP mice (EYFP, enhanced yellow fluorescent protein), all of the YFP$^+$ cells expressed CD3 or NK1.1, indicating that the Cre expression is restricted to NK, T, and NKT cells (fig. S9, A to B). In *Klra6*^{cre}DTA mice, there was a 50 to 75% decrease of Ly49$^+$CD8$^+$ T cells in the spleen and lymph nodes, whereas Ly49$^+$ NK cells did not show a significant reduction (fig. S9C), consistent with the preferential expression of Ly49F on CD8$^+$ T cells. *Klra6*^{cre}DTA mice did not appear to spontaneously develop any autoimmune disorders or exhibit changes in their frequencies of effector T cells, T follicular helper (T_{FH}) cells, or germinal center (GC) B cells up to 8 months of age (fig. S9, D to E). When mice were infected with lymphocytic choriomeningitis virus (LCMV)–Armstrong or influenza A-PR8 viruses, there was a surge of Ly49$^+$CD8$^+$ T cells in the blood of wild-type mice (Fig. 6, A and E), consistent with our previous observations of increased KIR$^+$CD8$^+$ T cells in patients with acute SARS-CoV-2 or influenza infection. However, the frequency of Ly49$^+$CD8$^+$ T cells in *Klra6*^{cre}DTA mice remained very low at all times after viral challenge, whereas Ly49$^+$ NK cells in *Klra6*^{cre}DTA mice showed only minor reductions compared with DTA (wild-type) mice (figs. S10A and S11A), indicating a selective and efficient ablation of Ly49$^+$CD8$^+$ T cells.

With either LCMV or influenza infection, *Klra6*^{cre}DTA mice showed no difference in viral clearance (figs. S10B and S11B) or levels of effector- and virus-specific CD4$^+$ and CD8$^+$ T cell responses (figs. S10C and S11C) compared with those of controls. However, *Klra6*^{cre}DTA mice, but not DTA mice, developed autoimmune pathology characterized by increased numbers of T_{FH} and GC B cells in the spleen (Fig. 6B) as well as glomerular nephritis (Fig. 6C) and immunoglobulin G (IgG) deposition (Fig. 6D) in the kidney 30 days after LCMV-Armstrong infection. Similarly, influenza-infected *Klra6*^{cre}DTA mice displayed more-severe inflammation and pathology characterized by peribronchial and interstitial accumulation of inflammatory cells, pulmonary consolidation, and fibrous tissue hyperplasia in the lung 60 days after influenza infection (Fig. 6E). Thus, this population of regulatory CD8$^+$ T cells appears to suppress autoimmunity that can develop after viral infection rather than having any discernable role in viral clearance.

Discussion

Here, we identify KIR$^+$CD8$^+$ T cells as an important regulatory CD8$^+$ T cell subset in humans. Previous studies have shown that KIR$^+$CD8$^+$ T cells are terminally differentiated cells and display a restricted TCR repertoire (*9*, *30*), which is consistent with our findings. Correlations between KIR$^+$CD8$^+$ T cells and tumor immune surveillance (*31*, *32*) or chronic viral infections (*33*, *34*) have been reported, but the suppressive functions of this population have not been clearly defined previously. We demonstrate the regulatory function of KIR$^+$CD8$^+$ T cells toward pathogenic CD4$^+$ T cells through an in vitro functional assay using PBMCs from CeD patients. This effect of KIR$^+$CD8$^+$ T cells seems specific to self-reactive or otherwise pathogenic T cells, but not CD4$^+$ T cells recognizing foreign antigens. Similar to the perforin- or Fas-FasL–dependent suppression of self-reactive CD4$^+$ T cells by murine Ly49$^+$CD8$^+$ T cells (*4*, *7*, *35*), human KIR$^+$CD8$^+$ T cells likely target pathogenic CD4$^+$ T cells through direct killing because KIR$^+$CD8$^+$ T cells significantly up-regulate cytotoxic molecules and suppress gliadin-specific CD4$^+$ T cells in a contact-dependent manner by inducing apoptosis. Additionally, the destruction of pathogenic CD4$^+$ T cells by KIR$^+$CD8$^+$ T cells appears to depend on recognition of both classical and nonclassical class I MHC molecules because the blockade of either HLA-ABC or HLA-E can reverse the suppression by KIR$^+$CD8$^+$ T cells. However, the MHC restrictions of the KIR$^+$ regulatory CD8$^+$ T cells at work vary between individuals. Because KIR receptors deliver inhibitory signals through their ITIMs to suppress the activation and functions of KIR$^+$CD8$^+$ T cells, antibody-dependent blockade of KIR3DL1 or KIR2DL3 may further enhance the suppressive activity of KIR$^+$CD8$^+$ T cells toward pathogenic CD4$^+$ T cells.

We frequently observed an increased frequency of KIR$^+$CD8$^+$ T cells in the blood as well as in the inflamed tissues of patients with autoimmune disease. This increase positively correlated with disease activity in CeD intestinal biopsies. The expansion of KIR$^+$CD8$^+$ T cells in the context of autoimmune diseases may act as a negative feedback mechanism to ameliorate pathogenesis by killing autoreactive T cells. Moreover, increased KIR$^+$CD8$^+$ T cells were found in SARS-CoV-2– or influenza-infected patients and were associated with

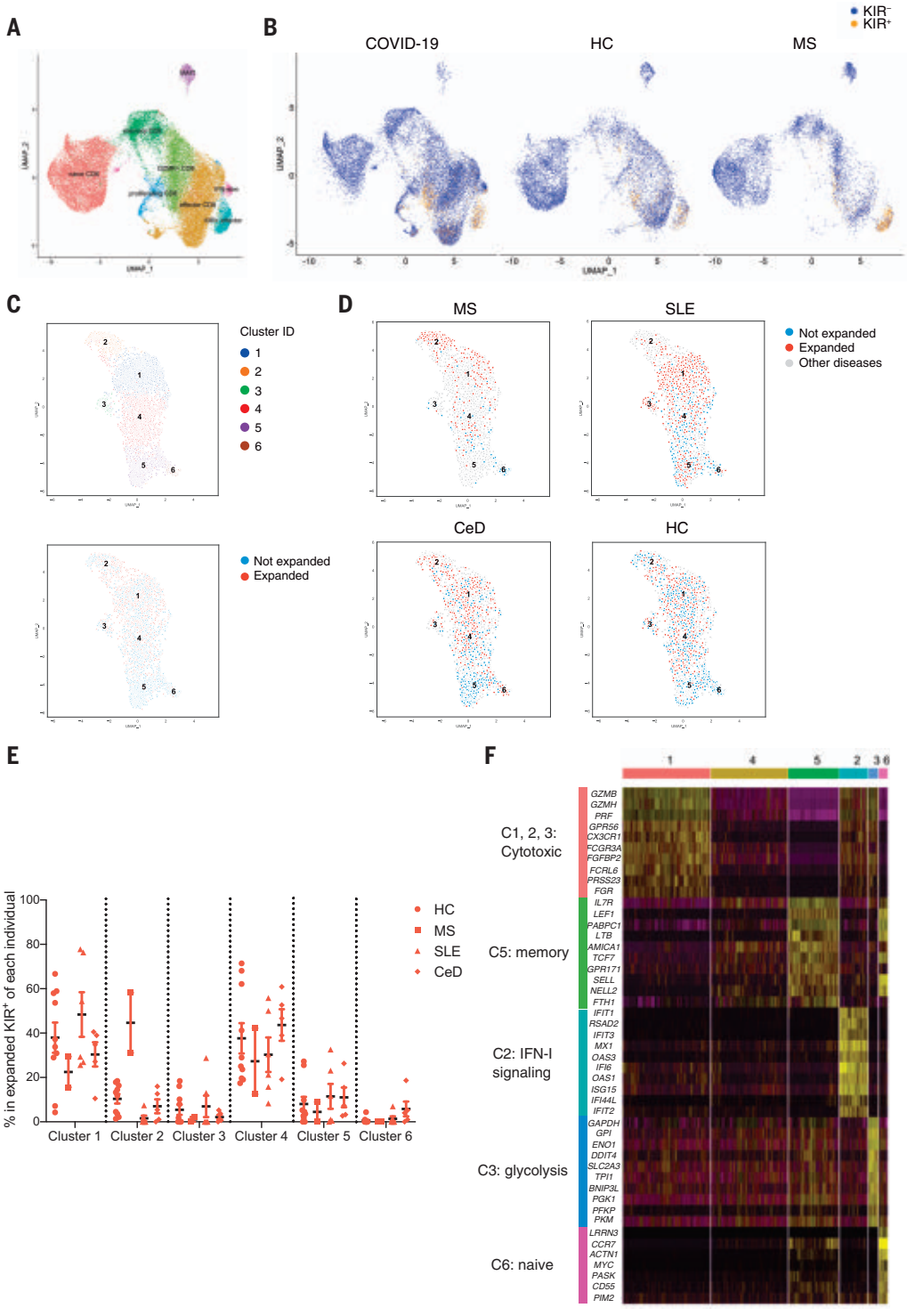


Fig. 4. scRNA-seq analysis of KIR$^+$CD8$^+$ T cells in the blood. (**A** and **B**) scRNA-seq analysis of total CD8$^+$ T cells from the blood of HCs (N = 10), MS patients (N = 6), and COVID-19 patients (N = 25) by 10X Genomics. (A) UMAP plot of the eight subpopulations identified by unsupervised clustering. (B) UMAP plots showing the distribution of KIR$^+$CD8$^+$ T cells (expressing *KIR3DL1*, *KIR3DL2*, *KIR2DL1*, or *KIR2DL3* transcripts) and KIR$^-$CD8$^+$ T cells from HCs, MS patients, and COVID-19 patients. (**C** to **F**) KIR$^+$CD8$^+$ T cells in the blood of HCs (N = 10) and patients with MS (N = 2), SLE (N = 6), and CeD (N = 5) were sorted for scRNA-seq using the Smart-seq2 protocol and analyzed using the R package "Seurat." (C) UMAP plots showing KIR$^+$CD8$^+$ T cells segregated into six clusters (top) and the distribution of expanded (≥2 cells expressing same TCR) and unexpanded (cells expressing unique TCRs) cells (bottom). (D) UMAP plots of KIR$^+$CD8$^+$ T cells from MS, SLE, and CeD patients and HCs are shown, with expanded and unexpanded cells annotated with different colors (expanded, red; unexpanded, blue; other diseases, gray). (E) Cluster compositions of expanded KIR$^+$CD8$^+$ T cells from each individual. (F) Heatmap showing expression of the top 10 genes differentially expressed in each cluster, with the categories of each group of genes annotated on the left.

autoimmune-related complications in COVID-19 patients. This suggests that increase of KIR$^+$CD8$^+$ T cells is a general mechanism induced during an infection, which, in classical murine studies, has been seen to break tolerance—that is, allow the activation of self-reactive T cells that normally require innate immune signals in addition to their cognate antigens (*36*). In mice, there was also a surge of circulating Ly49$^+$CD8$^+$ T cells after LCMV-Armstrong or influenza A-PR8 infection. Selective ablation of Ly49$^+$CD8$^+$ T cells did not interfere with antiviral immune responses but led to exacerbated autoimmunity after virus infection. This is in line with the autoimmune phenotypes secondary to LCMV infection in *Helios*$^{-/-}$ mice, in which both CD8$^+$ and CD4$^+$ regulatory T cells are defective (*6*). Therefore, we hypothesize that a major role of these CD8$^+$ regulatory T cells is to control autoreactive T cells that are activated during an infection, likely because they are cross-reactive to foreign antigens. This would allow an organism to have the benefit of a complete T cell repertoire while limiting damage from the autoreactive clones. This type of peripheral tolerance is distinct from and likely complementary to the one mediated

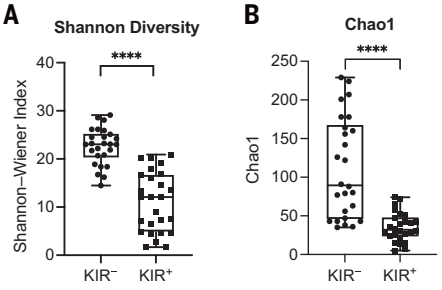


Fig. 5. Analysis of TCR sequences of KIR$^+$CD8$^+$ T cells. (**A** and **B**) Summary histograms showing Shannon-Wiener indices (A) and Chao estimates (B) of TCRs of KIR$^-$ versus KIR$^+$CD8$^+$ T cells from 26 subjects, including 11 healthy donors, two MS, five SLE, three CeD, and five T1D patients as evaluated by VDJtools. ****P < 0.0001; Wilcoxon matched-pairs signed-rank test.

by CD4$^+$ T_{regs}, which represent a separate lineage of T cells and do not appear to be generally active in the human infectious diseases analyzed here or in murine infections (*37*, *38*).

Although there is still much more to learn about KIR$^+$CD8$^+$ T cells and their murine equivalents, the data presented here and in previous studies indicate that they represent an important element in peripheral tolerance and in our understanding of the relationship between autoimmunity and infectious diseases. Further characterization of this pathway and how it may break down in autoimmune diseases and severe infections, like COVID-19, will be important challenges for the future. Likewise, our findings on the KIR$^+$CD8$^+$ T cells and their properties described here may be useful in understanding key cellular dynamics in immune dysregulation and in potential therapeutic approaches to suppress undesirable self-reactivity in autoimmune or infectious diseases.

Materials and methods

Human samples

Our study cohort of patients with autoimmune disorders met classification criteria for SLE (*39*), CeD (*40*), or MS (*41*), respectively. Collection of blood or biopsies from patients with SLE, CeD, or MS was covered by IRB-14734, IRB-20362, and IRB-36061. Blood samples from patients during influenza virus infection were obtained from patients who had influenza-like symptoms and were tested positive for influenza A virus at the Emergency Department or the Express Outpatient Clinic at Stanford Hospital, which is covered by IRB-22442. Blood from healthy subjects was requested from the Stanford Blood Center or drawn from healthy volunteers under IRB-40146. The protocols listed above have been approved by the Research Compliance Office of Stanford University. PBMCs from MS patients were also obtained from the Multiple Sclerosis Center at the University of California, San Francisco (UCSF), with the protocol approved by the committee on Human Research at UCSF. Informed written consent was obtained from all participants. Detailed information of the HCs and patients with autoimmune diseases included in the study is provided as table S5. PBMCs were isolated from the blood through density gradient centrifugation (Ficoll-Paque, GE Healthcare). Duodenal biopsies from CeD patients were treated twice with 6 mM EDTA in calcium/magnesium-free Hanks' balanced salt solution (HBSS) for 30 min at 37°C. Supernatants containing the epithelial fractions were combined, washed, and kept on ice until staining. The remaining tissues were minced and incubated with 200 μg/ml of Liberase TL and 20 U/ml of deoxyribonuclease (DNase) I in Iscove's modified Dulbecco's medium (IMDM) for 30 min at 37°C. Digested cell suspensions were passed through a 100-μm cell strainer, washed with complete media, and combined with the epithelial fraction for staining.

For sample collection of COVID-19 patients, enrollment included any adult with reverse transcription polymerase chain reaction (RT-PCR)–positive COVID-19. Informed consent was obtained from each patient or from the patient's legally authorized representative if the patient was unable to provide consent. Participants were excluded if they were taking any experimental medications (i.e., those medications not approved by a regulatory agency for use in COVID-19). COVID-19 severity of illness was defined as described in the literature (*42*). Collection of blood from COVID-19 patients was covered by IRB-55689 and NCT04373148. Handling of COVID-19 PBMCs for flow cytometric analysis was covered under APB-3343-MD0620. The IRB and APB protocols mentioned above have been approved by the Research Compliance Office of Stanford University. Clinical metadata were obtained from the Stanford clinical data electronic medical record system as per consented participant permission, and definitions and diagnoses of disease were used according to Harrison's Principles of Internal Medicine, 20e. Clinical metadata for the COVID-19 patients in this study are presented in table S6.

Mice

R26R-EYFP mice (stock no. 006148) and ROSA-DTA mice (stock no. 009669) were obtained from the Jackson Laboratory. *Klra6*^{cre} mice were generated by the Stanford Transgenic, Knockout and Tumor model Center. *Klra6* reporter mice were generated by crossing *Klra6*^{cre} mice to R26R-EYFP mice. *Klra6*^{cre}DTA mice were generated by crossing *Klra6*^{cre} mice to ROSA-DTA mice. ROSA-DTA (wild-type) littermates were used as controls in the experiments described here. All mice were housed in the specific pathogen-free animal facilities at Stanford University. Experiments in this study were covered by the following animal protocols approved by the Animal Care and Use Committee of Stanford University: APLAC-10081, APLAC-34021, and APLAC-32763.

Flow cytometric analysis

The fluorescent dye–conjugated antibodies used for staining were listed in table S7. Frozen cell samples were thawed and washed in 10% fetal bovine serum (FBS) with benzonase (Sigma-Aldrich, 25 U/ml) in RPMI. After 450*g* centrifugation, cells were treated with Fc receptor (FcR) block (Biolegend, 10 μg/ml) in FACS buffer [0.5% bovine serum albumin (BSA), 2 mM EDTA in phosphate-buffered saline (PBS)] for 10 min followed by staining with antibodies against surface molecules (30 min, 4°C). For intracellular staining, cells were fixed and permeabilized with the Intracellular Fixation & Permeabilization Buffer Set (eBioscience), followed by staining with antibodies against intracellular antigens (30 min, 4°C). Cells were acquired on an LSR II flow cytometer (BD), and data were analyzed using FlowJo X. Dead cells were excluded based on viability dye staining (LIVE/DEAD Fixable Near-IR Dead Cell Stain, ThermoFisher).

Functional assay

Chymotryptic gluten digests were deamidated with recombinant human transglutaminase 2, as described previously (*43*). PBMCs were isolated from blood of HLA-DQ2.5$^+$ CeD patients on day 0. CD8$^+$ T cells were purified from PBMCs using CD8 microbeads (Miltenyi) per manufacturer's instructions, stained with flow antibodies, and live CD3$^+$CD56$^-$CD8$^+$KIR$^+$ or KIR$^-$ T cells were sorted out by FACSAria Fusion flow cytometer (BD). The sorted KIR$^+$ and KIR$^-$CD8$^+$ T cells were stimulated with anti-CD3/CD28 beads (Gibco) at 1:1 ratio (1 μl of beads per 4×10^4 cells) supplemented with 50 U/ml of IL-2 in 96-well plates for 18 hours. KIR$^+$ and KIR$^-$ NK cells were sorted for PBMCs and rested overnight. The CD8$^-$ PBMCs were stimulated with 250 μg/ml of deamidated gluten or 10 μg/ml of influenza A HA 306-318 peptide (PKYVKQNTLKLAT) or left unstimulated at 3×10^5 to 1×10^6 cells/100 μl per well supplemented with 50 U/ml or 300 U/ml of IL-2. X-VIVO 15 with Gentamicin L-Gln (Lonza) supplemented with 10% human AB serum (Sigma-Aldrich) was used as culture medium. After 18 hours, anti-CD3/CD28 beads were removed from CD8$^+$ T cells by a magnet and KIR$^+$ or KIR$^-$CD8$^+$ T cells or NK cells were added to the culture of CD8$^-$ PBMCs at a 1:30 ratio. In the class I MHC blockade experiments, 10 μg/ml of anti–HLA-ABC (W6/32, Biolegend), anti–HLA-E (3D12, eBioscience),

Table 1. Selected specificity groups from the GLIPH2 analysis that contained TCRβ sequences from three or more individuals and exhibited significant bias of V-gene usage ($P < 0.05$). Significant motif residues are highlighted in red in CDR3 alignments.

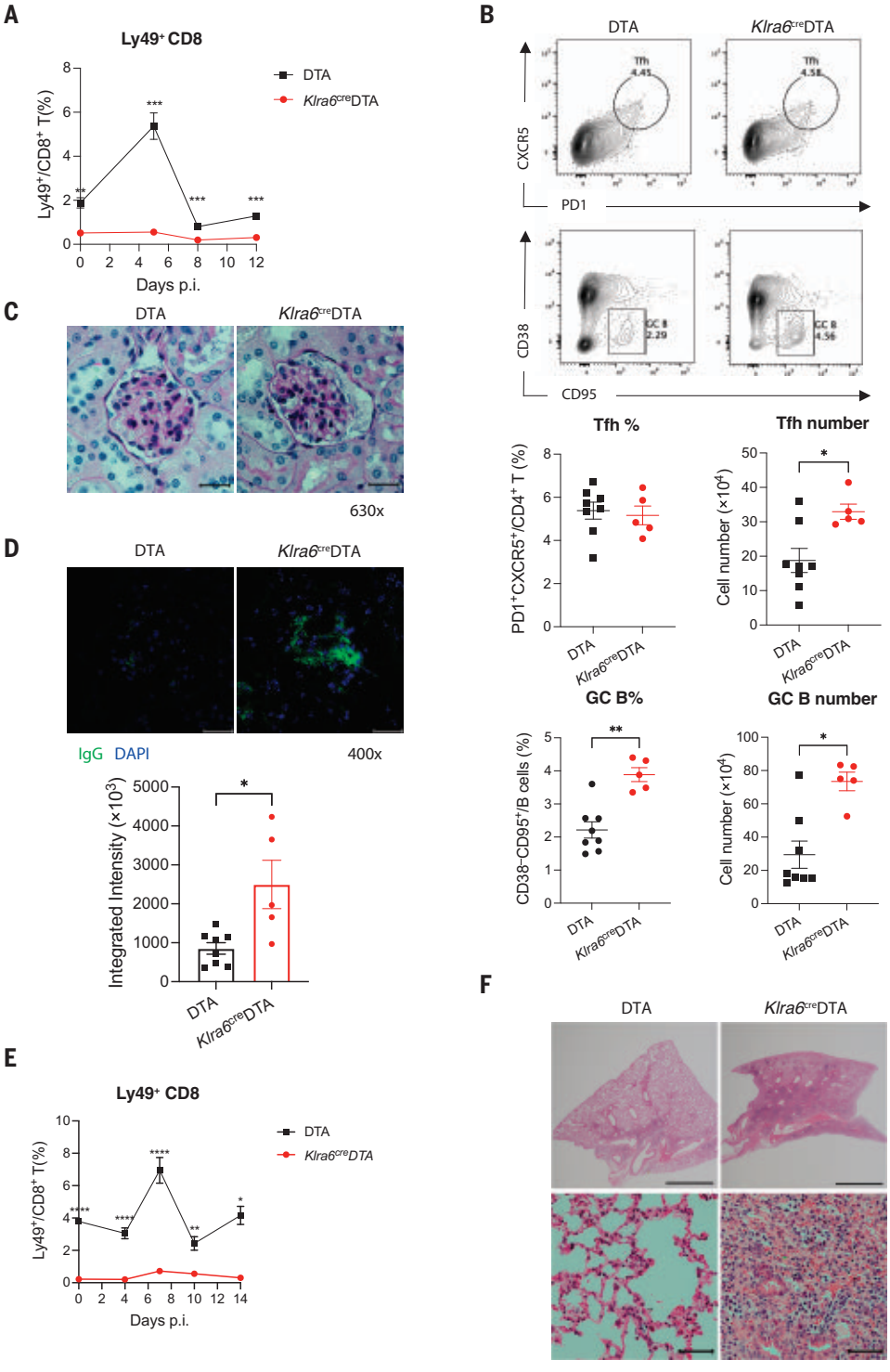
Group ID	Subject: Condition	CDR3β	TRBV	TRBJ	%	HLA-A	HLA-B	HLA-C
Group I motif-PFGE	CD229:CeD	CASGTGPFGELFF	TRBV10-2	TRBJ2-2	78.801	A*26:01/A*32:01	B*27:02/B*38:01!	C*02:02/C*12:03
	HC1505:HC	CASGTGPFGELFF	TRBV10-2	TRBJ2-2	54.545	A*11:01/A*34:01	B*38:02!/B*40:02	C*07:02/C*15:02!
	OHC018:COVID-19	CASSPFGETQYF	TRBV11-2	TRBJ2-5	0.427	A*24:02/A*29:02	B*08:01/B*51:01	C*07:01/C*15:02!
	0033:COVID-19	CASSPFGETQYF	TRBV25-1	TRBJ2-5	14.67	A*02:01/A*03:01	B*07:02/B*49:01	C*07:01/C*07:02
Group II global-S%DSGGTDT	T1D25:T1D	CASSFDSGGTDTQYF	TRBV7-2	TRBJ2-3	3.053	A*03:01!/A*03:01!	B*40:01/B*49:01!	C*03:04/C*07:01
	OHC015:COVID-19	CASSFDSGGTDTQYF	TRBV7-3	TRBJ2-3	0.629	A*02:01/A*11:02	B*27:04/B*35:01	C*03:03/C*12:02
	0033:COVID-19	CASSLDSGGTDTQYF	TRBV7-2	TRBJ2-3	0.016	A*02:01/A*03:01!	B*07:02/B*49:01!	C*07:01/C*07:02
Group III global-%EGLAGLQET	BBS660:SLE	CASSEGLAGLQETQYF	TRBV27	TRBJ2-5	7.286	A*24:02/A*25:01	B*08:01!/B*15:01	C*03:03/C*07:01
	CD224:CeD	CASSEGLAGLQETQYF	TRBV27	TRBJ2-5	0.541	A*01:01/A*02:01	B*08:01!/B*08:01!	C*07:01/C*07:01
	BBS660:SLE_KIR-	CASSEGLAGLQETQYF	TRBV27	TRBJ2-5	1.136	A*24:02/A*25:01	B*08:01!/B*15:01	C*03:03/C*07:01
	T1D51:T1D	CASGEGLAGLQETQYF	TRBV27	TRBJ2-5	0.313	A*02:01/A*23:17	B*08:01!/B*40:02	C*04:01/C*18:01
Group IV motif-SQDF	0033:COVID-19	CASSQDFWGRYEQYF	TRBV4-3	TRBJ2-7	0.016	A*02:01/A*03:01!	B*07:02/B*49:01	C*07:01/C*07:02
	KBSLE20:SLE	CASSQDFFNEQFF	TRBV4-1	TRBJ2-1	0.204	A*02:06/A*30:01	B*13:02!/B*15:25	C*04:03/C*06:02
	0033:COVID-19	CASSQDFEGGYEQYF	TRBV4-2	TRBJ2-7	0.016	A*02:01/A*03:01!	B*07:02/B*49:01	C*07:01/C*07:02
	HC0231:HC	CASSQDFRGLSSRSYNEQFF	TRBV3-1	TRBJ2-1	0.004	A*03:01!/A*03:01!	B*35:03/B*52:01	C*12:02/C*12:03
	KBSLE20:SLE	CASSQDFLNEQFF	TRBV4-1	TRBJ2-1	0.815	A*02:06/A*30:01	B*13:02!/B*15:25	C*04:03/C*06:02
	KBSLE20:SLE	CASSQDFYNEQFF	TRBV4-1	TRBJ2-1	0.407	A*02:06/A*30:01	B*13:02!/B*15:25	C*04:03/C*06:02
	HC4084:HC	CASSQDFGGIYGYTF	TRBV14-1	TRBJ1-2	0.061	A*03:02!/A*11:01!	B*35:01/B*35:03	C*04:01/C*04:01
	HC2061:HC_KIR-	CASSQDFSRPYGYTF	TRBV4-3	TRBJ1-2	2.439	A*11:01!/A*11:01!	B*27:05/B*55:01	C*01:02/C*03:03
	KBSLE20:SLE	CASSQDFRNEQYF	TRBV4-1	TRBJ2-7	0.407	A*02:06/A*30:01	B*13:02!/B*15:25	C*04:03/C*06:02
	KBSLE20:SLE	CASSQDFHNEQFF	TRBV4-1	TRBJ2-1	0.204	A*02:06/A*30:01	B*13:02!/B*15:25	C*04:03/C*06:02
	KBSLE20:SLE	CASSQDFGGEQFF	TRBV4-1	TRBJ2-1	0.407	A*02:06/A*30:01	B*13:02!/B*15:25	C*04:03/C*06:02
	KBSLE07:SLE	CASSQDFRGGNTEAFF	TRBV3-1	TRBJ1-1	0.214	A*02:03/A*11:01!	B*13:01!/B*18:01	C*03:04/C*07:04
	HC0230:HC	CASSQDFPGQEETQYF	TRBV4-2	TRBJ2-5	0.787	A*02:06/A*33:03	B*13:01!/B*58:01	C*03:02/C*03:04
	KBSLE20:SLE	CASSQDFNNEQFF	TRBV4-1	TRBJ2-1	1.629	A*02:06/A*30:01	B*13:02!/B*15:25	C*04:03/C*06:02
	HC1505:HC	CASSQDFLLQGDLEEQYF	TRBV3-1	TRBJ2-7	0.023	A*11:01!/A*34:01	B*38:02/B*40:02	C*07:02/C*15:02
	HC0231:HC	CASSQDFSYNSPLHF	TRBV3-1	TRBJ1-6	0.011	A*03:01!/A*03:01!	B*35:03/B*52:01	C*12:02/C*12:03
Group V motif-PFF	KBSLE22:SLE	CASSPFFGDTEAFF	TRBV6-5	TRBJ1-1	0.84	A*33:03!/A*66:02	B*44:03!/B*58:01!	C*03:02/C*07:01
	BBS706:SLE	CASSPFFPGGPYEQYF	TRBV6-5	TRBJ2-7	22.609	A*02:01/A*02:01	B*44:02!/B*44:02!	C*05:01/C*05:01
	BBS706:SLE	CASSPFFPGDPYEQYF	TRBV6-5	TRBJ2-7	0.29	A*02:01/A*02:01	B*44:02!/B*44:02!	C*05:01/C*05:01
	HC0230:HC	CASSFPFFAEAFF	TRBV11-3	TRBJ1-1	0.11	A*02:06/A*33:03!	B*13:01/B*58:01!	C*03:02/C*03:04
Group IV motif-SQEI	OHC014:COVID-19	CASSQEIRGWDTEAFF	TRBV3-1	TRBJ1-1	0.011	A*11:01/A*23:01	B*35:02/B*49:01!	C*04:01/C*07:01
	HC1505:HC	CASSQEIMTAYGYTF	TRBV4-1	TRBJ1-2	0.023	A*11:01/A*34:01	B*38:02/B*40:02	C*07:02/C*15:02
	BBS706:SLE_KIR-	CASSQEIAGSSYNEQFF	TRBV3-1	TRBJ2-1	0.242	A*02:01/A*02:01	B*44:02/B*44:02	C*05:01!/C*05:01!
	0033:COVID-19	CASSQEISQPQHF	TRBV4-1	TRBJ1-5	0.016	A*02:01/A*03:01	B*07:02/B*49:01!	C*07:01/C*07:02
	HC2419:HC	CASSQEIGRATDTQYF	TRBV4-2	TRBJ2-3	0.215	A*02:06/A*24:02	B*40:01/B*48:01	C*08:01/C*12:03
	KBSLE18:SLE	CASSQEIPITGNTIYF	TRBV4-1	TRBJ1-3	0.238	A*30:01/A*32:01	B*15:01/B*44:02	C*04:01/C*05:01!
	HC2164:HC	CASSQEIAGGETEAFF	TRBV4-2	TRBJ1-1	0.006	A*02:01/A*24:02	B*15:13/B*44:02	C*05:01!/C*08:01
	0033:COVID-19	CASSQEIGVVETQYF	TRBV3-1	TRBJ2-5	0.423	A*02:01/A*03:01	B*07:02/B*49:01!	C*07:01/C*07:02
	HC0231:HC	CASSQEIGQLNQPQHF	TRBV4-1	TRBJ1-5	0.004	A*03:01/A*03:01	B*35:03/B*52:01	C*12:02/C*12:03

or isotype controls were added to the culture. In the transwell assay, $CD8^-$ PBMCs were cultured with 250 μg/ml of deamidated gluten and 50 U/ml of IL-2 in the lower chamber, and preactivated KIR^+CD8^+ T cells were added to the upper chamber of the Millicell-96 Cell Culture Insert Plate (Millipore Sigma, cat. no. PSHT004R1). On day 3, 50 U/ml of IL-2 was added to the cultures. Cells were harvested on day 6 and stained with 10 μg/ml of HLA-DQ2.5 tetramers (*44*) complexed with four disease-relevant and immunodominant gliadin T cell epitopes (DQ2.5-glia-α1a, QLQPFPQPELPY; DQ2.5-glia-α2, PQPELPYPQPE; DQ2.5-glia-ω1, QQPFPQPEQPFP; and DQ2.5-glia-ω2, FPQPEQPFPWQP) (*45*) or 10 μg/ml of HLA-DR4 tetramers complexed with influenza A HA 306-318 peptide for 45 min at room temperature. Magnetic bead enrichment of $tetramer^+$ $CD4^+$ T cells was performed as previously described (*46*). Cells were washed with FACS buffer and then stained with antibodies against surface molecules for 30 min at 4°C. After two washes with FACS buffer, 10% of the cells were reserved for FACS analysis and 90% were labeled with anti–phycoerythrin (PE) microbeads and subjected to magnetic bead enrichment of PE-conjugated tetramer-positive cells using a single magnetic activated cell sorting (MACS) column according to the manufacturer's protocol (Miltenyi). Cells were also harvested on day 3 to measure annexin V binding (BD) on gliadin-specific $CD4^+$ T cells. All cells were acquired on an LSR II flow cytometer (BD), gated on live $CD3^+CD4^+CD8^-TCR\alpha\beta^+$ cells, and analyzed using FlowJo X software. The frequency of tetramer-positive cells was calculated by dividing the number of post-enrichment $tetramer^+$ $CD4^+$ T cells by the number of $CD4^+$ T cells in the pre-enrichment sample multiplied by 9 (to account for the fact that 90% of the cells were used for the enrichment).

Bulk RNA-seq gene-expression quantification and data analysis

Live KIR^+ or KIR^- $CD8^+$ T cells were bulk sorted directly into 500 μl of TRIzol (ThermoFisher, cat. no.15596026) by FACSAria Fusion flow cytometer (BD). Total RNA was extracted from TRIzol samples using chloroform separation

Fig. 6. Exacerbated autoimmunity in *Klra6*^{cre}DTA mice after viral infection. (**A**) Frequency of Ly49$^+$CD8$^+$ T cells in the blood of DTA mice (N = 8) and *Klra6*^{cre}DTA mice (N = 5) 0, 5, 8, and 12 days after LCMV-Armstrong infection. $^{**}P < 0.01$; $^{***}P < 0.001$; repeated measures (RM) two-way analysis of variance (ANOVA) followed by multiple comparisons test. p.i., postinfection. (**B**) Representative contour plots and summarized scatter plots displaying frequency and absolute number of PD1$^+$CXCR5$^+$CD4$^+$ T cells (T_{FH}) and CD38$^-$CD95$^+$ GC B cells in the spleen of DTA mice (N = 8) and *Klra6*^{cre}DTA mice (N = 5) 30 days after LCMV-Armstrong infection. $^*P < 0.05$; $^{**}P < 0.01$; Mann-Whitney test. Tfh, T follicular helper cell. (**C**) Representative kidney pathology of day-30 LCMV-infected DTA mice and *Klra6*^{cre}DTA mice assessed by PAS staining (630X; scale bars, 20 μm). (**D**) IgG deposition in glomeruli of the kidneys of DTA mice (N = 8) and *Klra6*^{cre}DTA mice (N = 5) 30 days after LCMV infection accessed by immunofluorescence staining (400X; scale bars, 20 μm) and quantified by ImageJ. $^*P < 0.05$; Mann-Whitney test. DAPI, 4′,6-diamidino-2-phenylindole. (**E**) Frequency of Ly49$^+$CD8$^+$ T cells in the blood of DTA mice (N = 6) and *Klra6*^{cre}DTA mice (N = 5) after influenza A-PR8 infection. $^*P < 0.05$; $^{**}P < 0.01$; $^{****}P < 0.0001$; RM two-way ANOVA followed by multiple comparisons test. (**F**) Microscopy of representative H&E staining of lung sections from DTA mice and *Klra6*^{cre}DTA mice 60 days after influenza infection (top, 20X; scale bars, 2 mm) (bottom, 630X; scale bars, 50 μm). Representative data from two independent experiments are shown. The means ± SEMs are indicated.



and isopropanol precipitation and then RNAeasy Plus Mini kit (Qiagen) for clean-up. After analysis on the 2100 Bioanalyzer (Agilent), the sequencing libraries were prepared using the SMARTer Stranded RNA-seq Kit (Clontech). The resulting library was sequenced on the HiSeq 4000 platform (Illumina) in Stanford Functional Genomics Facility. For each sample in the bulk RNA-seq library, 75–base pair (bp) paired-end reads were acquired from the sequencer. We aligned the reads to the human reference genome (NCBI GRCh38) using STAR v2.7.0e (*47*). Gene counts were quantified and normalized (TPM) with Salmon (*48*). DEGs were determined using the DESeq function (adjusted $P < 0.05$, fold change > 2) in the DESeq2 R package (*49*). Heatmaps were generated with seaborn.clustermap in python. Gene Ontology analysis plots were generated with the R package "clusterProfiler." To generate gene sets for GSEA, we selected the top 200 genes upregulated in Ly49$^+$CD8$^+$ T cells compared with Ly49$^-$CD8$^+$ T cells in EAE mice (*7*) and the previously reported CD4$^+$ T_{reg} signature genes identified in mice (*17*). These mouse genes were converted to homolog genes in humans and constituted as gene sets for the subsequent

GSEA analysis (*15*, *16*) in human KIR$^+$ versus KIR$^-$CD8$^+$ T cells.

scRNA-seq analysis of kidneys and synovial tissues

The unique molecular identifier (UMI) count matrices of cells in kidneys (accession code SDY997) (*50*) or synovial tissues (accession code SDY998) (*51*) generated by CEL-Seq2 were downloaded from the ImmPort repository, and downstream analysis was performed using the Seurat 3.0 package (*52*). Cells with fewer than 1000 detected genes, more than 5000 detected genes, or more than 25% mitochondrial genes were discarded. CD8$^+$ T cells (expressing *CD3D*, *CD3E*, *CD8A*, and *CD8B* transcripts) and CD4$^+$ T cells (expressing *CD3D*, *CD3E*, and *CD4* transcripts) were selected for standard downstream procedures of log-normalization, variable gene selection, and data scaling.

scRNA-seq analysis of bronchoalveolar immune cells

Filtered expression matrix of scRNA-seq of immune cells from the bronchoalveolar lavage fluid of six severe and three moderate COVID-19 patients and three HCs generated by 10X Genomics (*23*) were downloaded from Gene Expression Omnibus under the accession number GSE145926. CD8$^+$ T cells were identified for downstream analysis using the Seurat 3.0 package (*52*).

scRNA-seq analysis of blood CD8$^+$ T cells by 10X Genomics

scRNA-seq of T cells from the blood of healthy subjects (N = 10), MS patients (N = 6), and COVID-19 patients (N = 25; ArrayExpress: E-MTAB-9357) (*24*) from the microfluidic droplet platform (10X Genomics Chromium Single Cell 5′-paired-end chemistry) were demultiplexed, aligned to the GRCh38 reference genome, and converted into gene counts matrices using CellRanger 3.1.0. Downstream analysis was performed using the Seurat 3.0 package (*52*). Cells with fewer than 800 detected genes, more than 3000 detected genes, or more than 10% mitochondrial genes were discarded. CD8$^+$ T cells (expressing *CD8A* and *CD8B* but not *TRDC* transcripts) were selected for further analysis. To make counts comparable among cells, gene counts were normalized to 10,000 reads per cell, then log-transformed. We identified highly variable genes for each individual, then integrated gene expression data from all individuals using Seurat's integration anchor discovery algorithm (*53*). We performed principal components analysis (PCA) dimensionality reduction on the integrated data, then clustered cells with the Louvain algorithm and visualized the data using uniform manifold approximation and projection (UMAP). We identified canonical cell type marker genes that were conserved across conditions using the Wilcoxon rank-sum test implemented in the Seurat package's "FindConservedMarkers" function.

Quantification of scRNA-seq gene expression by Smart-seq2 and data analysis

Blood KIR$^+$CD8$^+$ T cells (live CD3$^+$CD56$^-$CD8$^+$TCRαβ$^+$KIR$^+$ cells) were sorted into 96-well plates, and cDNA synthesis was performed using the Smart-seq2 protocol (*26*) with minor modifications described previously (*54*). cDNA products were purified with 0.65X AMPure XP beads (Beckman Coulter) on the Biomek FX Automated Workstation (Beckman Coulter) and eluted with 25 μl of water. Then 2 μl of the purified products were subjected to quality control using capillary electrophoresis on a Fragment Analyzer (Agilent Technologies) by Stanford Protein and Nucleic Acid Facility.

cDNA in 96-well plates was transferred into 384-well low volume serial dilution (LVSD) plates (TTP Labtech) and diluted to 160 ng/ml using a Mosquito X1 liquid handler (TTP Labtech). Illumina sequencing libraries were prepared as described previously (*55*) using a Mosquito HTS liquid handler (TTP Labtech). After library preparation, wells of each library plate were pooled using a Mosquito HTS liquid handler (TTP Labtech). Pooling was followed by two purifications using 0.65X and 1X AMPure XP beads (Beckman Coulter), respectively. Library quality was assessed by Agilent 2100 Bioanalyzer and normalized to 5 nM. Libraries were sequenced on the Hiseq4000 Sequencing System (Illumina) in Stanford Functional Genomics Facility, acquiring 150-bp paired-end reads. FASTQ files for each cell were extracted and generated, distinguished by the unique dual index adapters. Reads were aligned to the GRCh38 genome using STAR v2.6.1d. Transcript abundance was quantified using HTSeq v0.5.4p5.

Standard procedures for filtering, log-normalization, variable gene selection, dimensionality reduction, and clustering were performed using the Seurat 3.0 package (*52*). Briefly, cells with fewer than 800 detected genes, more than 5000 detected genes, or more than 15% mitochondrial genes were discarded. To make counts comparable among cells, gene counts were normalized to 10,000 reads per cell, then log-transformed. After PCA dimensionality reduction, cells were clustered by running the Louvain algorithm and visualized using UMAP. Differential expression analysis was performed using the Wilcoxon rank-sum test implemented in the Seurat package's "FindAllMarkers" function. Significant DEGs were defined as those with log fold change >0.5 and Bonferroni-corrected $P < 0.05$.

Single-cell TCR-seq

TCR sequencing (TCR-seq) was performed using our previously developed single-cell paired TCR-seq method (*27*). For the first TCR reaction, 1 μl of the cDNA products of Smart-seq2 was preamplified with HotStarTaq DNA polymerase (Qiagen) using multiplex PCR with multiple Vα and Vβ region primers, Cα and Cβ region primers. Three steps of PCR were performed followed by purification of 350- to 380-bp products using a Qiaquick gel extraction kit (Qiagen). The purified product was then sequenced on a Miseq platform (Illumina) acquiring 2 × 250 bp reads.

Bulk TCRβ sequencing

KIR$^+$CD8$^+$ T cells were sorted from PBMCs of nine healthy subjects, and DNA was extracted using QIAamp DNA Micro Kit (Qiagen). Sequencing of the CDR3 regions of human TCRβ chains was performed using the immunoSEQ Assay by Adaptive Biotechnologies.

GLIPH2 analysis

Single-cell TCR sequences of sorted KIR$^+$ (6815 unique TCRs) and KIR$^-$CD8$^+$ T cells (1630 unique TCRs) from HCs (N = 10), MS (N = 2), SLE (N = 20), CeD (N = 5), T1D (N = 5), and COVID-19 (N = 5) patients and bulk TCRβ sequences of sorted KIR$^+$CD8$^+$ T cells from nine HCs (5607 unique TCRβ) along with their class I HLA alleles were used as inputs. The GLIPH2 analysis generated 982 clusters, and 668 of them were shared between any two sources. We further filtered the resulting GLIPH clusters to 62 specificity groups that contained TCRs from three or more individuals and exhibited significant bias of V-gene usage ($P < 0.05$), and some of them are shown in Table 1.

In vitro cell proliferation assay

CD8$^+$ T cells were purified from PBMCs of healthy donors using CD8 microbeads (Miltenyi) per manufacturer's instructions, stained with flow antibodies, and live CD3$^+$CD56$^-$CD8$^+$KIR$^+$ or KIR$^-$ T cells were sorted out by FACSAria Fusion flow cytometer (BD). The sorted KIR$^+$ or KIR$^-$CD8$^+$ T cells were stimulated with anti-CD3/CD28 beads (Gibco) at 1:1 ratio (1 μl beads per 4 × 10^4 cells) supplemented with 50 U/ml of IL-2 in 96-well plates for 18 hours. The CD8$^-$ PBMCs were labeled with CellTrace Violet (CTV, ThermoFisher) per manufacturer's instruction. 1 μg/ml of anti-CD3 (UCHT-1) was coated on 96-well plate in 50 μl of PBS per well at 4°C overnight. After removal of anti-CD3/CD28 microbeads, KIR$^+$ and KIR$^-$ CD8$^+$ T cells were mixed with CTV-labeled CD8$^-$ PBMCs at 1:30, 1:15, or 1:10 ratios and cultured in 96-well plate precoated with anti-CD3. After 3 days, cells were harvested and dilution of CTV in CD4$^+$ T cells was analyzed by flow cytometry.

LCMV-Armstrong infection

Two-month-old female *Klra6*^{cre}DTA mice and ROSA-DTA (wild-type) littermates were infected

with 2×10^5 PFU LCMV-Armstrong through intraperitoneal injection. On day 30, mice were euthanized, and spleen lymphocytes were subjected to flow cytometric analysis. After kidneys were harvested, one of them was fixed with formalin and embedded in paraffin, and the other was embedded in optimal cutting temperature (OCT) compound and flash frozen.

Murine influenza virus infection

Influenza A/PR/8/34 H1N1 was obtained from Charles River (cat. no. 10100374, lot no. 4XP201023), aliquoted and stored in −80°C. A new aliquot was thawed and diluted with sterile PBS for infection each time. Twenty-week-old male *Klra6*^{cre}DTA mice and ROSA-DTA littermates were anesthetized and intratracheally infected with a sublethal dose of influenza A/PR/8/34 H1N1 in 20 μl of PBS. Mice were euthanized on day 60 and lungs were harvested, fixed with formalin, and embedded in paraffin.

Immunohistochemistry

To evaluate the glomerular nephritis in the kidneys of LCMV-infected mice, tissue sections were generated from formalin-fixed paraffin-embedded (FFPE) kidneys and Periodic acid–Schiff (PAS) stain was performed by Stanford Animal Histology Services. For analysis of IgG deposition in kidney, frozen sections from mouse kidneys were stained with Alexa Fluor 488 Goat anti-Mouse IgG (H+L) (Invitrogen), and images were acquired using Leica SP8. Quantification of fluorescence intensity was performed using the ImageJ software. To assess immunopathology in the lungs of influenza-infected mice, tissue sections were generated from FFPE lung tissues and stained with hematoxylin and eosin (H&E).

Quantification of viral load by RT-qPCR

RNA was extracted from the blood of LCMV-infected mice or the lungs of influenza-infected mice using RNeasy Mini Kit (Qiagen) and converted to cDNA using Applied Biosystems High-Capacity cDNA Reverse Transcription Kit (cat. no. 43-688-14). The generated cDNA was then amplified by quantitative PCR (qPCR) using Platinum SYBR Green qPCR SuperMix-UDG w/ROX (cat. no. 11744100) per manufacturer's instructions with the following primers: Beta Actin primer 1: 5′-CGA GGC CCA GAG CAA GAG AG-3′, Beta Actin primer 2: 5′-CGG TTG GCC TTA GGG TTC AG-3′, LCMV GP forward primer: 5′-TGC CTG ACC AAA TGG ATG ATT-3′, LCMV GP reverse primer: 5′-CTG CTG TGT TCC CGA AAC ACT-3′, NS1 forward primer: 5′-TGT CAA GCT TTC AGG TAG ATT G-3′, NS-1 reverse primer: 5′-CTC TTA GGG ATT TCT GAT CTC-3′, M1 forward primer: 5′-AAG ACC AAT CCT GTC ACC TCT GA-3′, and M1 reverse primer: 5′-CAA AGC GTC TAC GCT GCA GTC C-3′.

Statistical analysis

No specific statistical methods were used to predetermine sample size. No samples or data points were excluded from the analysis. All results are presented as the means ± SEMs. The significance of the difference between groups was analyzed as described in the figure legends. P values <0.05 were considered statistically significant. All statistical analyses were performed using GraphPad Prism Software version 9.1.0.

REFERENCES AND NOTES

1. H. Nakagawa, L. Wang, H. Cantor, H. J. Kim, New Insights Into the Biology of CD8 Regulatory T Cells. *Adv. Immunol.* **140**, 1–20 (2018). doi: 10.1016/bs.ai.2018.09.001; pmid: 30366517
2. D. R. Koh *et al.*, Less mortality but more relapses in experimental allergic encephalomyelitis in CD8$^{-/-}$ mice. *Science* **256**, 1210–1213 (1992). doi: 10.1126/science.256.5060.1210; pmid: 1589800
3. H. Jiang, S. I. Zhang, B. Pernis, Role of CD8$^+$ T cells in murine experimental allergic encephalomyelitis. *Science* **256**, 1213–1215 (1992). doi: 10.1126/science.256.5060.1213; pmid: 1375398
4. H. J. Kim, B. Verbinnen, X. Tang, L. Lu, H. Cantor, Inhibition of follicular T-helper cells by CD8$^+$ regulatory T cells is essential for self tolerance. *Nature* **467**, 328–332 (2010). doi: 10.1038/nature09370; pmid: 20844537
5. H. J. Kim *et al.*, CD8$^+$ T regulatory cells express the Ly49 Class I MHC receptor and are defective in autoimmune prone B6-Yaa mice. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2010–2015 (2011). doi: 10.1073/pnas.1018974108; pmid: 21233417
6. H. J. Kim *et al.*, Stable inhibitory activity of regulatory T cells requires the transcription factor Helios. *Science* **350**, 334–339 (2015). doi: 10.1126/science.aad0616; pmid: 26472910
7. N. Saligrama *et al.*, Opposing T cell responses in experimental autoimmune encephalomyelitis. *Nature* **572**, 481–487 (2019). doi: 10.1038/s41586-019-1467-x; pmid: 31391585
8. M. M. Rahim, A. P. Makrigiannis, Ly49 receptors: Evolution, genetic diversity, and impact on immunity. *Immunol. Rev.* **267**, 137–147 (2015). doi: 10.1111/imr.12318; pmid: 26284475
9. N. K. Björkström *et al.*, CD8 T cells express randomly selected KIRs with distinct specificities compared with NK cells. *Blood* **120**, 3455–3465 (2012). doi: 10.1182/blood-2012-03-416867; pmid: 22968455
10. E. Estefanía *et al.*, Human KIR2DL5 is an inhibitory receptor expressed on the surface of NK and T lymphocyte subsets. *J. Immunol.* **178**, 4402–4410 (2007). doi: 10.4049/jimmunol.178.7.4402; pmid: 17371997
11. V. Abadie, L. M. Sollid, L. B. Barreiro, B. Jabri, Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu. Rev. Immunol.* **29**, 493–525 (2011). doi: 10.1146/annurev-immunol-040210-092915; pmid: 21219178
12. H. Sjöström *et al.*, Identification of a gliadin T-cell epitope in coeliac disease: General importance of gliadin deamidation for intestinal T-cell recognition. *Scand. J. Immunol.* **48**, 111–115 (1998). doi: 10.1046/j.1365-3083.1998.00397.x; pmid: 9716100
13. H. Sheng *et al.*, Distinct PLZF$^+$CD8$\alpha\alpha^+$ Unconventional T Cells Enriched in Liver Use a Cytotoxic Mechanism to Limit Autoimmunity. *J. Immunol.* **203**, 2150–2162 (2019). doi: 10.4049/jimmunol.1900832; pmid: 31554695
14. Y. Vuddamalay, J. P. van Meerwijk, CD28$^-$ and CD28^{low}CD8$^+$ Regulatory T Cells: Of Mice and Men. *Front. Immunol.* **8**, 31 (2017). doi: 10.3389/fimmu.2017.00031; pmid: 28167946
15. A. Subramanian *et al.*, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005). doi: 10.1073/pnas.0506580102; pmid: 16199517
16. V. K. Mootha *et al.*, PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003). doi: 10.1038/ng1180; pmid: 12808457
17. D. Zemmour *et al.*, Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nat. Immunol.* **19**, 291–301 (2018). doi: 10.1038/s41590-018-0051-0; pmid: 29434354
18. W. Yu *et al.*, Clonal Deletion Prunes but Does Not Eliminate Self-Specific αβ CD8$^+$ T Lymphocytes. *Immunity* **42**, 929–941 (2015). doi: 10.1016/j.immuni.2015.05.001; pmid: 25992863
19. L. F. Su, B. A. Kidd, A. Han, J. J. Kotzin, M. M. Davis, Virus-specific CD4$^+$ memory-phenotype T cells are abundant in unexposed adults. *Immunity* **38**, 373–383 (2013). doi: 10.1016/j.immuni.2012.10.021; pmid: 23395677
20. L. Quintana-Murci, Human Immunology through the Lens of Evolutionary Genetics. *Cell* **177**, 184–199 (2019). doi: 10.1016/j.cell.2019.02.033; pmid: 30901539
21. M. M. Davis, P. Brodin, Rebooting Human Immunology. *Annu. Rev. Immunol.* **36**, 843–864 (2018). doi: 10.1146/annurev-immunol-042617-053206; pmid: 29490162
22. P. S. Ohashi *et al.*, Ablation of "tolerance" and induction of diabetes by virus infection in viral antigen transgenic mice. *Cell* **65**, 305–317 (1991). doi: 10.1016/0092-8674(91)90164-T; pmid: 1901764
23. M. Liao *et al.*, Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020). doi: 10.1038/s41591-020-0901-9; pmid: 32398875
24. Y. Su *et al.*, Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19. *Cell* **183**, 1479–1495. e20 (2020). doi: 10.1016/j.cell.2020.10.037; pmid: 33171100
25. G. X. Zheng *et al.*, Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017). doi: 10.1038/ncomms14049; pmid: 28091601
26. S. Picelli *et al.*, Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014). doi: 10.1038/nprot.2014.006; pmid: 24385147
27. A. Han, J. Glanville, L. Hansmann, M. M. Davis, Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* **32**, 684–692 (2014). doi: 10.1038/nbt.2938; pmid: 24952902
28. H. Huang, C. Wang, F. Rubelt, T. J. Scriba, M. M. Davis, Analyzing the *Mycobacterium tuberculosis* immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* **38**, 1194–1202 (2020). doi: 10.1038/s41587-020-0505-4; pmid: 32341563
29. I. Tassi *et al.*, Phospholipase C-γ2 is a critical signaling mediator for murine NK cell activating receptors. *J. Immunol.* **175**, 749–754 (2005). doi: 10.4049/jimmunol.175.2.749; pmid: 16002670
30. S. A. Schattgen *et al.*, Integrating T cell receptor sequences and transcriptional profiles by clonotype neighbor graph analysis (CoNGA). *Nat. Biotechnol.* **40**, 54–63 (2022). doi: 10.1038/s41587-021-00989-2; pmid: 34426704
31. L. Gimeno *et al.*, KIR+ CD8+ T Lymphocytes in Cancer Immunosurveillance and Patient Survival: Gene Expression Profiling. *Cancers* **12**, 2991 (2020). doi: 10.3390/cancers12102991; pmid: 33076479
32. S. Chakraborty *et al.*, Providence of the CD25$^+$KIR$^+$CD127$^-$FOXP3$^-$CD8$^+$ T-cell subset determines the dynamics of tumor immune surveillance. *Immunol. Cell Biol.* **96**, 1035–1048 (2018). doi: 10.1111/imcb.12166; pmid: 29768737
33. W. Lu *et al.*, Suppression of HIV Replication by CD8$^+$ Regulatory T-Cells in Elite Controllers. *Front. Immunol.* **7**, 134 (2016). doi: 10.3389/fimmu.2016.00134; pmid: 27148256
34. G. Alter *et al.*, Ligand-independent exhaustion of killer immunoglobulin-like receptor-positive CD8$^+$ T cells in human immunodeficiency virus type 1 infection. *J. Virol.* **82**, 9668–9677 (2008). doi: 10.1128/JVI.00341-08; pmid: 18579582
35. K. Akane, S. Kojima, T. W. Mak, H. Shiku, H. Suzuki, CD8$^+$CD122$^+$CD49d^{low} regulatory T cells maintain T-cell homeostasis by killing activated T cells via Fas/FasL-mediated cytotoxicity. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2460–2465 (2016). doi: 10.1073/pnas.1525098113; pmid: 26869716
36. Q. Yin *et al.*, Nanoparticle-enabled innate immune stimulation activates endogenous tumor-infiltrating T cells with broad antigen specificities. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2016168118 (2021). doi: 10.1073/pnas.2016168118; pmid: 34021082
37. R. J. Betts, A. W. Ho, D. M. Kemeny, Partial depletion of natural CD4$^+$CD25$^+$ regulatory T cells with anti-CD25 antibody does not alter the course of acute influenza A virus infection. *PLOS ONE* **6**, e27849 (2011). doi: 10.1371/journal.pone.0027849; pmid: 22125630
38. M. C. Rogers *et al.*, CD4$^+$ Regulatory T Cells Exert Differential Functions during Early and Late Stages of the Immune Response to Respiratory Viruses. *J. Immunol.* **201**, 1253–1266 (2018). doi: 10.4049/jimmunol.1800096; pmid: 29997123
39. M. C. Hochberg, Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* **40**, 1725 (1997). doi: 10.1002/art.1780400928; pmid: 9324032
40. J. F. Ludvigsson *et al.*, Diagnosis and management of adult coeliac disease: Guidelines from the British Society of Gastroenterology. *Gut* **63**, 1210–1228 (2014). doi: 10.1136/gutjnl-2013-306578; pmid: 24917550

41. C. H. Polman *et al*., Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann. Neurol.* **69**, 292–302 (2011). doi: 10.1002/ana.22366; pmid: 21387374
42. G. Chen *et al*., Clinical and immunological features of severe and moderate coronavirus disease 2019. *J. Clin. Invest.* **130**, 2620–2629 (2020). doi: 10.1172/JCI137244; pmid: 32217835
43. V. Abadie *et al*., IL-15, gluten and HLA-DQ8 drive tissue destruction in coeliac disease. *Nature* **578**, 600–604 (2020). doi: 10.1038/s41586-020-2003-8; pmid: 32051586
44. H. Quarsten *et al*., Staining of celiac disease-relevant T cells by peptide-DQ2 multimers. *J. Immunol.* **167**, 4861–4868 (2001). doi: 10.4049/jimmunol.167.9.4861; pmid: 11673490
45. L. M. Sollid, S. W. Qiao, R. P. Anderson, C. Gianfrani, F. Koning, Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics* **64**, 455–460 (2012). doi: 10.1007/s00251-012-0599-z; pmid: 22322673
46. T. J. Scriba *et al*., Ultrasensitive detection and phenotyping of CD4$^+$ T cells with optimized HLA class II tetramer staining. *J. Immunol.* **175**, 6334–6343 (2005). doi: 10.4049/jimmunol.175.10.6334; pmid: 16272285
47. A. Dobin *et al*., STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi: 10.1093/bioinformatics/bts635; pmid: 23104886
48. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017). doi: 10.1038/nmeth.4197; pmid: 28263959
49. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). doi: 10.1186/s13059-014-0550-8; pmid: 25516281
50. A. Arazi *et al*., The immune cell landscape in kidneys of patients with lupus nephritis. *Nat. Immunol.* **20**, 902–914 (2019). doi: 10.1038/s41590-019-0398-x; pmid: 31209404
51. F. Zhang *et al*., Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat. Immunol.* **20**, 928–942 (2019). doi: 10.1038/s41590-019-0378-1; pmid: 31061532
52. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018). doi: 10.1038/nbt.4096; pmid: 29608179
53. T. Stuart *et al*., Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e21 (2019). doi: 10.1016/j.cell.2019.05.031; pmid: 31178118
54. H. Huang *et al*., Select sequencing of clonally expanded CD8$^+$ T cells reveals limits to clonal expansion. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8995–9001 (2019). doi: 10.1073/pnas.1902649116; pmid: 30992377
55. Tabula Muris Consortium, Overall coordination, Logistical coordination, Organ collection and processing, Library preparation and sequencing, Computational data analysis, Cell type annotation, Writing group, Supplemental text writing group, Principal investigators, Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature* **562**, 367–372 (2018). doi: 10.1038/s41586-018-0590-4; pmid: 30283141
56. Davis Lab, davis-lab-stanford/kir-cd8: 2022-01 - v1.0.0, version 1.0.0, Zenodo (2022); https://doi.org/10.5281/zenodo.5835290.

ACKNOWLEDGMENTS

We thank members of the Davis laboratory for helpful discussions. We also thank J. Coller from the Stanford Functional Genomics Facility for next-generation sequencing (NGS) (NIH award S10OD018220); R. Sinha, S. D. Conley, and I. L. Weissman for assistance in scRNA-seq library preparation; L. Chen for help with Python scripts; D. Chen from the Institute for Systems Biology for help with the COVID-19 10X scRNA-seq data processing; R. Wittman and R. Puri from the Stanford occupational health clinic; and I. Chang, J. Krempski, E. Do, M. Manohar, A. Fernandez, and W. Zhang for collecting and processing blood samples from COVID-19 patients. We thank Z. He and J. Fitzpatrick for data management and entry of COVID-19 patients; D. Furman, B. Gaudillière, and D. Feyaerts for their help in obtaining COVID-19 PBMCs; and J. J. Goronzy for providing the LCMV-Armstrong strain. We thank the Stanford Protein and Nucleic Acid Facility for performing fragment analysis; the Stanford Shared FACS facility for assistance in flow cytometric analysis and cell sorting; the Stanford Animal Histology Services for the histology; the Stanford Transgenic, Knockout and Tumor model Center for assistance in generating *Klra6*^{cre} mice; and the Stanford Cell Sciences Imaging Facility for assistance in microscopy. **Funding:** We are grateful for funding by the National Institute of Allergy and Infectious Diseases (NIAID) (U19-AI057229) and the Howard Hughes Medical Institute to M.M.D. The collection of MS samples was supported by a grant from the National Multiple Sclerosis Society (NMSS), RG-1611-26299. This work was partially supported by pilot funding from the Stanford Diabetes Research Center (P30DK116074). The project described was also supported in part by award no. 1S10OD010580-01A1 from the National Center for Research Resources (NCRR). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NCRR or the National Institutes of Health (NIH). M.Z. was supported by a National Science Foundation Graduate Research Fellowship. J.R.H. and Y.S. were funded by the Wilke Family Foundation, Merck, and the Biomedical Advanced Research and Development Authority under contract no. HHSO102016000031C. L.M.St. and M.J.S. were supported by NIAID under award no. R01AI139550. S.L.H. is supported by the NIH/NINDS (R35NS111644) and the Valhalla Foundation. C.L.D. received funding from NIAID under award nos. U19-AI057229 and U19-AI090019. C.K. was funded by NIH R01DK063158. P.J.U. was supported by NIAID (R01AI125197-04) and the Henry Gustav Floren Trust. K.C.N. was funded by the Sunshine Foundation and the Sean N. Parker Center at Stanford University. N.S. was supported by a postdoctoral fellowship and a Career Transition grant from NMSS. **Author contributions:** Jin.L., N.S., and M.M.D. conceived the study. Jin.L. and N.S. performed the experiments. M.Z., Jin.L., S.-H.C., J.P., and A.T.S. analyzed the scRNA-seq data. M.J.S., L.C., and L.M.St. contributed to analysis of bulk RNA-seq data. X.J. prepared the bulk RNA-seq libraries. A.C. and L.M.So. provided the HLA-DQ2.5 molecules. Jie.L. prepared samples for bulk TCRβ sequencing. J.W. and A.M.M. performed single-cell TCR-seq. Y.S. and J.R.H. generated the scRNA-seq data of COVID-19 patients. V.V.A.M. prepared the HLA-DR4 molecules. B.A.P. and C.K. provided deamidated gluten. L.B.K., J.E.D., S.L.H., and J.R.O. provided PBMCs from MS patients. K.B., W.H.R., and P.J.U. provided PBMCs from SLE patients. K.C.N. and N.Q.F.-B. provided blood and biopsy samples from CeD patients. K.C.N. and G.K.R.D. provided PBMCs and clinical metadata of COVID-19 patients. C.L.D. provided PBMCs from influenza-infected patients. Jin.L. and M.M.D. wrote the manuscript with helpful input from all authors. **Competing interests:** N.S., M.M.D., and Jin.L. are coinventors on patent application 62/882,810, which includes discoveries described in this manuscript. S.L.H. serves on the board of directors for Neurona; serves on scientific advisory boards for Accure, Alector, Annexon, and Molecular Stethoscope; and has received travel reimbursement and writing assistance from F. Hoffmann-La Roche and Novartis for CD20-related meetings and presentations. M.M.D. is involved in a start-up company, Mozart Therapeutics, Inc., for the clinical applications of these discoveries. L.M.So. has served as a consultant for Mozart Therapeutics, Inc. J.E.D. has received honoraria for consulting services within the last three years from Alexion, Biogen, Bristol Myers Squibb, Celgene, Genentech, Genzyme, Janssen Pharmaceuticals, and Novartis. J.R.H. is a PACT Pharma Board Member for Isoplexis Corp and is on the board for Research Funding from Merck and Gilead. The other authors declare that they have no competing interests. **Data and materials availability:** Raw and processed RNA-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus under the accession number GSE193442. All custom-made code can be accessed at https://github.com/davis-lab-stanford/kir-cd8 and has been deposited at Zenodo (56). Biotinylated HLA-DQ molecules are available under a material transfer agreement with the University of Oslo from L.M.So. All other data are available in the main text or the supplementary materials. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abi9591
Figs. S1 to S11
Tables S1 to S7
MDAR Reproducibility Checklist

12 April 2021; resubmitted 12 October 2021
Accepted 1 March 2022
Published online 8 March 2022
10.1126/science.abi9591

Pushing the Boundaries of Knowledge

As AAAS's first multidisciplinary, open access journal, *Science Advances* publishes research that reflects the selectivity of high impact, innovative research you expect from the *Science* family of journals, published in an open access format to serve a vast and growing global audience. Check out the latest findings or learn how to submit your research: **ScienceAdvances.org**

Science Advances
AAAS

GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

Compartmentalized dendritic plasticity during associative learning

Simon d'Aquin, Andras Szonyi, Mathias Mahn, Sabine Krabbe, Jan Gründemann, Andreas Lüthi*

INTRODUCTION: Learning is mediated by experience-dependent plasticity of neuronal function. Although we have a detailed knowledge of synaptic and dendritic plasticity in vitro, learning-induced functional changes in vivo are mostly assessed by use of soma-centric methods such as unit recordings or imaging of somatic Ca^{2+} activity. However, these methods do not reveal the complexity of dendritic signaling and its regulation by local neuronal circuitry. Moreover, because dendrites are functionally independent subcellular compartments that dynamically integrate incoming information and thereby affect a neuron's input-output function, the questions arise whether compartmentalized plasticity occurs during learning and which mechanisms ultimately control somatic output in behaving animals.

RATIONALE: To investigate dendritic function and plasticity in vivo, we focused on the lateral amygdala (LA), a subcortical brain structure that is central to classical auditory fear conditioning, a fast and robust form of associative learning. During auditory fear conditioning, an auditory conditioned stimulus (CS; typically a pure tone or white noise) is paired with an aversive unconditioned stimulus (US; typically a foot shock), which results in the induction of Hebbian activity-dependent synaptic potentiation at auditory synaptic inputs onto LA principal neurons (PNs). This view has recently been extended by studies reporting that similar proportions of neurons up- and down-regulate their CS response upon fear conditioning, suggesting that fear learning involves more diverse forms of plasticity. However, these studies relied on measurements of somatic activity, whereas dendritic activity and plasticity during associative learning was not explored.

RESULTS: To image the activity of dendrites and somas of amygdala PNs deep in the brain of awake mice undergoing classical fear conditioning, we used gradient-index lens–based high-resolution two-photon microscopy across multiple days. We show that sensory stimulation induces compartmentalized dendritic responses controlled by dendrite-targeting somatostatin-expressing (SST+) interneurons. Spontaneous inputs to PN dendrites are suppressed by SST+ interneurons, whereas salient sensory stimuli transiently alleviate SST+ interneuron–mediated inhibition of PN dendrites likely through VIP+ interneurons. In most cases, this evokes highly correlated somatic and dendritic sensory responses. However, sensory input can also lead to isolated dendritic responses without concomitant somatic output, indicating that dendrites of LA PNs can integrate auditory inputs locally.

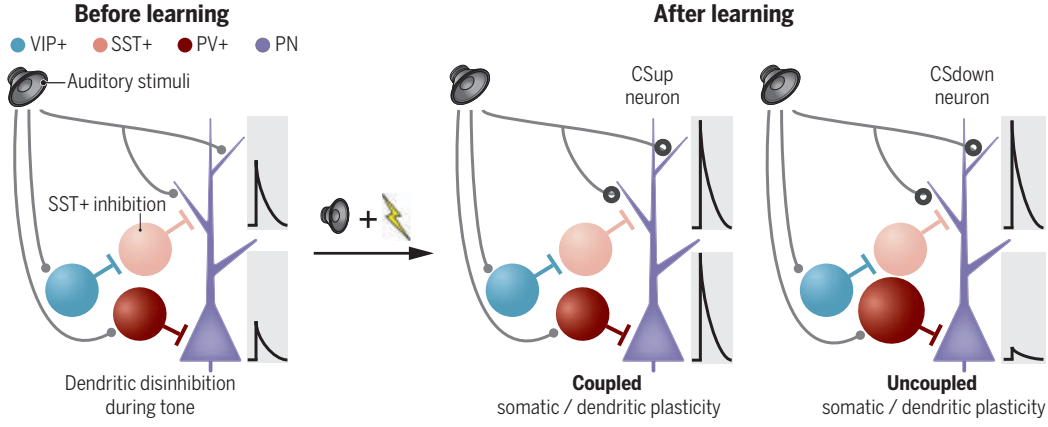
The relief of SST+ interneuron–mediated dendritic inhibition is necessary to amplify dendritic CS responses during conditioning, which is consistent with the notion that disinhibition through SST+ interneurons opens a temporal window during which CS inputs are eligible for the induction of associative dendritic plasticity upon concomitant exposure to an aversive US.

Fear conditioning induces bidirectional plasticity of somatic CS responses that correlates with learning at the behavioral level. On average, dendritic CS responses increase in neurons with both up-regulated (CSup) or down-regulated (CSdown) somatic CS responses. However, fear conditioning also increases the variance of CS responses across the dendritic tree of a given neuron, indicating that not all dendrites are undergoing similar levels of plasticity during learning. Moreover, dendritic spines that show up-regulated CS responses are more likely to be located on dendrites that exhibit increased CS responses after learning. Last, even though dendritic CS responses of CSdown neurons are potentiated, their somatic responses are reduced by enhanced Parvalbumin-expressing (PV+) interneuron–mediated perisomatic inhibition, counteracting the learning-induced increased synaptic drive.

CONCLUSION: Our findings demonstrate that LA PNs locally integrate dendritic sensory inputs in a compartmentalized manner and that fear conditioning–induced plasticity of dendritic and somatic sensory responses can be uncoupled. Both compartmentalized dendritic integration and uncoupling of dendritic and somatic plasticity are regulated by local inhibitory circuits that specifically target dendrites or the perisomatic region, two distinct subcellular compartments of postsynaptic PNs. The regulation of compartmentalized dendritic and somatic plasticity increases the computational capacity of amygdala circuits, possibly enhancing an animal's behavioral flexibility in the face of danger. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: andreas.luthi@fmi.ch
Cite this article as S. d'Aquin *et al.*, *Science* **376**, eabf7052 (2022); DOI: 10.1126/science.abf7052

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf7052



Fear conditioning induces compartmentalized plasticity. (**Left**) During auditory stimuli, dendritic disinhibition through SST+ interneurons amplifies dendritic Ca^{2+} responses. (**Right**) Fear conditioning induces a correlated increase of somatic and dendritic CS responses in a population of LA PNs (CSup neurons), whereas CSdown neurons show increased dendritic but decreased somatic CS responses. This compartmentalized plasticity is mediated by soma-targeting PV+ interneurons.

RESEARCH ARTICLE

NEUROSCIENCE

Compartmentalized dendritic plasticity during associative learning

Simon d'Aquin[1,2]†, Andras Szonyi[1,3], Mathias Mahn[1], Sabine Krabbe[1]‡, Jan Gründemann[1,4]‡, Andreas Lüthi[1,2]*

Experience-dependent changes in behavior are mediated by long-term functional modifications in brain circuits. Activity-dependent plasticity of synaptic input is a major underlying cellular process. Although we have a detailed understanding of synaptic and dendritic plasticity in vitro, little is known about the functional and plastic properties of active dendrites in behaving animals. Using deep brain two-photon Ca^{2+} imaging, we investigated how sensory responses in amygdala principal neurons develop upon classical fear conditioning, a form of associative learning. Fear conditioning induced differential plasticity in dendrites and somas regulated by compartment-specific inhibition. Our results indicate that learning-induced plasticity can be uncoupled between soma and dendrites, reflecting distinct synaptic and microcircuit-level mechanisms that increase the computational capacity of amygdala circuits.

Dendrites are active neuronal compartments that dynamically integrate synaptic inputs and thereby affect a neuron's input-output function (*1*, *2*). In vivo studies have shown that dendritic Ca^{2+} activity can be tuned to ongoing task variables (*3*, *4*), in some cases even restricted to distinct dendritic branches (*5*–*9*). However, a large proportion of Ca^{2+} activity in behaving animals occurs in the form of global transients that involve both the somas and dendrites of single neurons, indicating a high degree of functional coupling between these compartments and emphasizing the importance of recording activity from dendrites and their parent somas simultaneously (*10*–*12*).

Recent studies focused mainly on dendritic activity in cortical regions, whereas the functional properties of dendrites in subcortical brain areas remain unexplored. The lateral amygdala (LA) is a subcortical brain structure central to classical auditory fear conditioning (FC), a fast and robust form of associative learning (*13*). During auditory FC, a tone conditioned stimulus (CS) is paired with an aversive unconditioned stimulus (US; typically a foot shock), which results in the induction of Hebbian activity-dependent synaptic plasticity at auditory synaptic inputs onto LA principal neurons (PNs) (*14*–*18*). A similar proportion of neurons up- and down-regulate their CS response during learning (*19*–*21*). In addition, both plastic populations contained US responsive as well as US nonresponsive neurons, suggesting that FC involves more diverse forms of plasticity (*22*). However, these studies relied on somatic activity as readouts, whereas local dendritic activity and plasticity during associative learning was not explored.

Results

Deep brain imaging of tone and shock responses in LA neuron somas and dendrites in awake mice

To monitor somatic and dendritic activity in LA PNs, we virally expressed a Cre-dependent version of the Ca^{2+} sensor GCaMP6s together with highly diluted adeno-associated virus (AAV) expressing Cre recombinase under the calcium- and calmodulin-dependent protein kinase II (CaMKII) promotor and implanted a gradient-index (GRIN) lens above the virus injection site in the LA (Fig. 1, A to C; fig. S1A; and table S2). This resulted in sparse labeling of LA PNs with minimal background signal from out-of-focus sources when imaged with a two-photon microscope (Fig. 1, E and F, and movies S1 and S2). Somas and dendrites were registered and could be identified across multiple imaging sessions in awake mice, allowing reliable tracking of their activity across days (Fig. 1D).

The LA lacks a laminar organization, which makes challenging the association of dendritic branches with their parent soma. We therefore reconstructed the dendritic arbors of imaged neurons around the imaging plane on the basis of a three-dimensional (3D) structural scan acquired after completion of the behavioral experiments (Fig. 1G and fig. S2). To study associative fear learning, tones and mild electrical shocks are commonly used as CS and US, respectively. Our approach allowed us to reliably record spontaneous activity as well as CS and US responses in somas and different dendritic compartments of individual, reconstructed LA PNs in awake mice (Fig. 1, H and I).

Global and local spontaneous somatic and dendritic activity

To investigate the extent of functional coupling between somas and dendrites, we detected Ca^{2+} transients in identified dendritic branches of single neurons and compared them with somatic transients (Fig. 2, A and C). The amplitude of simultaneous somatic and dendritic transients was correlated (Fig. 2D) (*10*, *11*). Many of these dendritic transients may reflect back-propagating action potentials (bAPs) (Discussion). However, the strength of this correlation was reduced with increasing distance from the soma (Fig. 2E), indicating that activity in distal dendrites is more decorrelated from the soma. Overall, a small proportion of all dendritic Ca^{2+} transients in LA PNs (18.9% of 5227 transients detected in 113 dendrites from 13 neurons in six mice) were exclusive to dendritic branches, with no coincident Ca^{2+} event at the soma. Accordingly, the amplitude distribution of the somatic Ca^{2+} signals normalized to the maximal somatic signal magnitude and time locked to dendritic transients exhibited a bimodal distribution, with a peak around zero that reflects local dendrite-only transients and a second peak that reflects correlated somatic and dendritic transients (Fig. 2, F and G). The amplitude distribution of dendritic transients did not follow such a bimodal distribution (Fig. 2, H and I). The number and fraction of simultaneously active dendrites in a given cell was larger when dendrites were co-active with the soma as compared with when the soma was not active (Fig. 2, J to O).

Sensory stimulation generates local dendritic activity regulated by inhibition

We next examined soma-dendrite coupling during sensory stimulation in naïve animals. To check for differential sensory tuning in somas and dendrites, we presented mice with tones of varying frequencies and intensities as well as with electrical stimulation of different intensities (fig. S3, A to D). Although individual somas and dendrites exhibited a wide range of tuning properties, on average LA PNs responded stronger to high-intensity CS and US presentations compared with the stimuli with low intensity (fig. S3, B to D). Moreover, dendrites had similar tuning as that of their parent soma (fig. S3, E to G). However, the rate of dendrite-only Ca^{2+} transients was larger during tone presentations as compared with spontaneous activity in the absence of auditory input (Fig. 2P). Dendrite-specific Ca^{2+}

[1]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. [2]University of Basel, Basel, Switzerland. [3]Laboratory of Cellular Neurophysiology, Institute of Experimental Medicine, Budapest, Hungary. [4]Department of Biomedicine, University of Basel, Basel, Switzerland.
*Corresponding author. Email: andreas.luthi@fmi.ch
†Present address: Institute of Pharmacology and Toxicology, University of Zürich, Zürich, Switzerland.
‡Present address: Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Bonn, Germany.

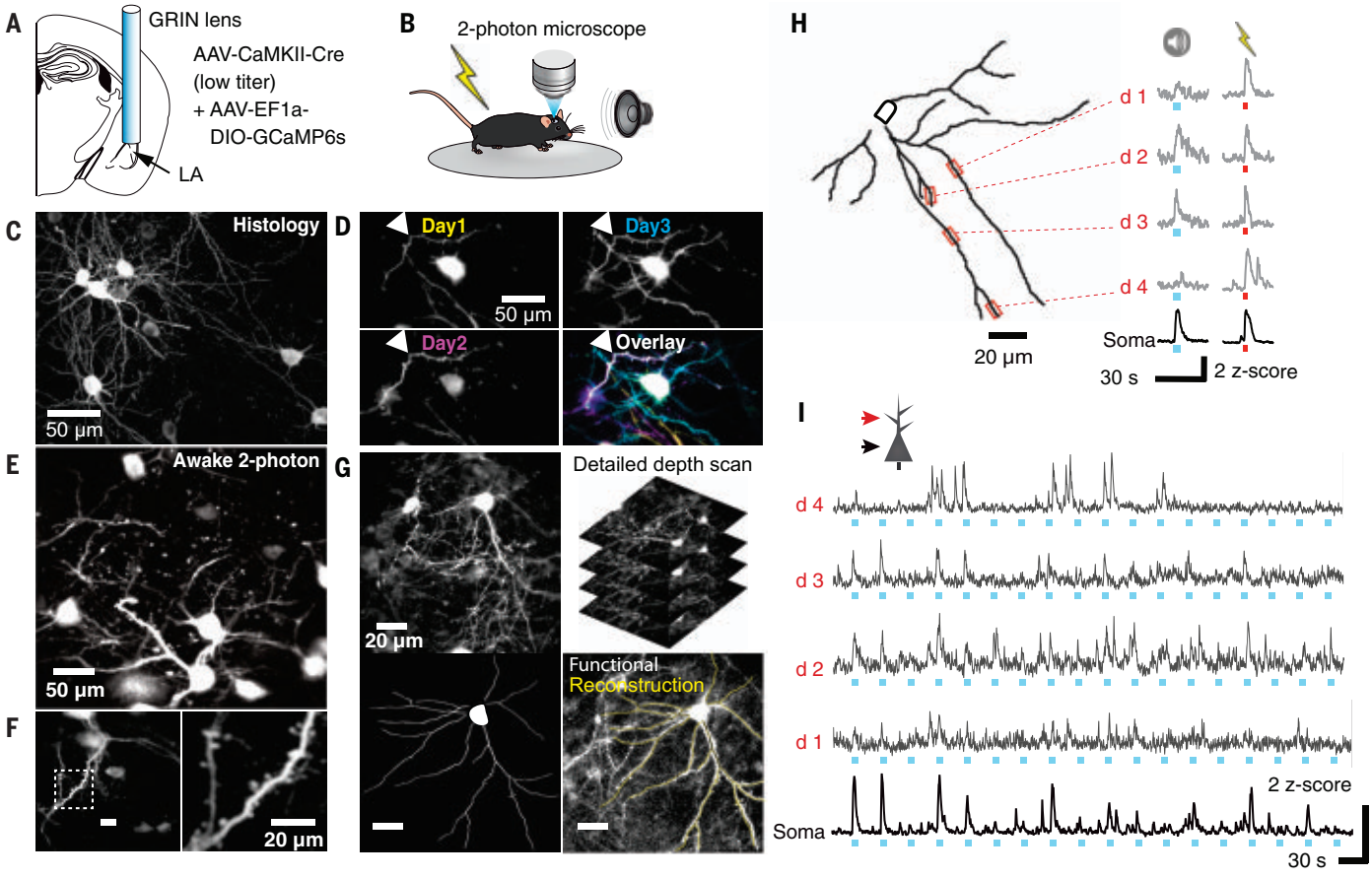


Fig. 1. In vivo dendritic calcium imaging in the LA. Experimental approach. (**A**) A highly diluted viral vector encoding CaMKII-Cre allows sparse but robust expression of GCaMP6s in LA PNs. A GRIN lens is implanted to gain optical access to the LA. (**B**) Head-fixed mice are allowed to run freely on a wheel under a two-photon microscope while presented with tone and shock stimuli. (**C**) Confocal image of sparsely infected LA PNs expressing GCaMP6s. (**D**) Imaging plane relocalization in an awake mouse over 3 consecutive days. Each image was obtained by averaging 50 imaging frames recorded at 30 Hz. The white arrowhead indicates the same dendritic segment active over days. (Bottom right) Overlay of the 3 days, one color per day. (**E**) Maximum intensity projection over a 15-min two-photon imaging session. (**F**) Deep brain imaging with dendritic spine resolution through a GRIN lens. The image was obtained by averaging 50 imaging frames recorded at 30 Hz. (**G**) Partial reconstruction of the dendritic arbor of imaged neurons. The dendrites of the imaged neurons were (bottom left) tracked in three dimensions on the basis of (top row) a structural scan acquired under anesthesia after the experiment and (bottom right) mapped back onto the functional imaging data. (**H**) Reconstruction of part of a LA PN dendritic tree. (**I**) Simultaneous Ca^{2+} imaging in the soma and dendrites of a LA PN. Shown is the same neuron as in (H). Blue rectangles indicate tones.

transients, but not dendritic Ca^{2+} transients coincidentally detected with somatic events, were stronger during tone presentations (Fig. 2, Q to T).

Similar to neocortical pyramidal cells, the activity of LA PNs is tightly controlled by local inhibitory interneurons. Somatostatin-positive (SST+) interneurons form γ-aminobutyric acid (GABA)–releasing synapses onto PN dendrites and are inhibited during CS and US presentations (*23*, *24*), which suggests that SST+ interneurons can gate dendritic excitability of LA PNs during sensory stimulation (*25*). We thus reduced SST+ interneuron output by virally expressing the inhibitory DREADD (designer receptor exclusively activated by designer drugs) hM4D(Gi) while imaging LA PN activity during tone presentations (Fig. 3, A to C).

Chemogenetic inhibition of SST+ interneurons increased the rate and proportion of dendrite-specific Ca^{2+} transients (Fig. 3, D to F) but had no further effect on tone-induced dendrite-specific Ca^{2+} transients, indicating that inhibition of SST+ interneurons leads to an increase in spontaneous, weak inputs that would otherwise be suppressed. Inhibition of SST+ interneurons led to an overall decrease in the amplitude correlation of dendritic and somatic transients (Fig. 3G). The rate and amplitude of somatic Ca^{2+} transients remained largely unaffected, with only a slight reduction in the somatic response integral (Fig. 3, H to J, and fig. S4, A to L).

To address whether CS-induced inhibition of SST+ interneurons allows for the generation of dendritic CS responses during learning, we chemogenetically enhanced SST+ interneuron output during CS-US pairing while imaging LA PNs (Fig. 3K). In animals expressing hM3D(Gq) in SST+ interneurons, clozapine N-oxide (CNO) administration during conditioning led to a reduction in dendritic CS response magnitude relative to CS responses measured in LA PNs during the preceding habituation session (fig. S4, N to T). This effect was even more pronounced when selectively analyzing PN dendrites and somas that showed CS responses during the habituation session (Fig. 3, L to O). Application of CNO in control animals only expressing mCherry did not reduce CS response magnitude (Fig. 3, P to T).

Associative FC induces bidirectional plasticity of somatic CS responses

One-photon Ca^{2+} imaging experiments indicate that FC induces a bidirectional plasticity of somatic CS responses in BLA PNs and that CS responses can be potentiated even in cells that do not exhibit US responses at the soma (*19*, *21*). Assuming that somatic plasticity reflects, to a large extent, plasticity at the synaptic and dendritic level, we sought to investigate FC-induced changes in CS responses at the subcellular level.

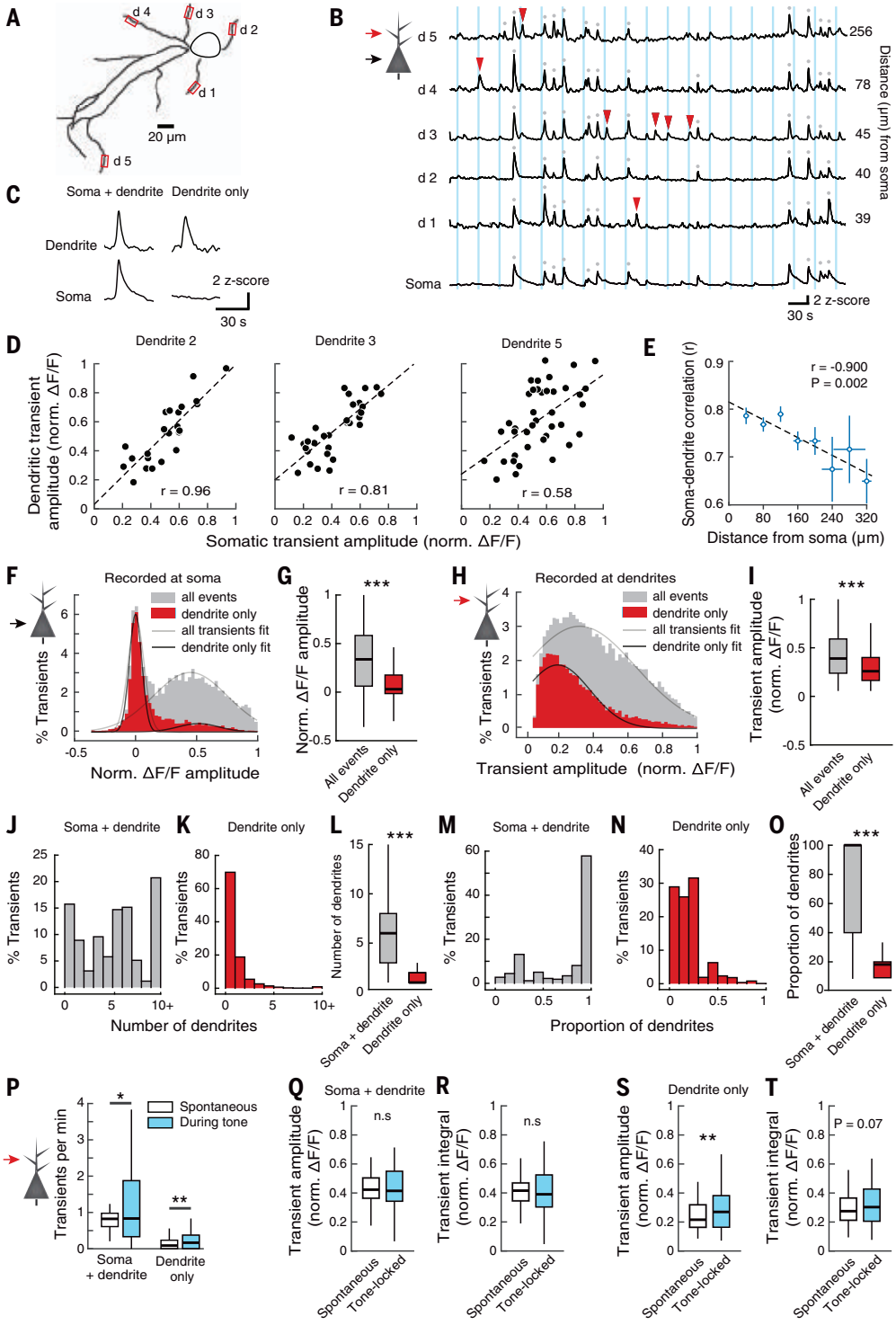


Fig. 2. Increased dendrite-specific activity during auditory stimulation. (**A**) Reconstructed dendritic arbor of a recorded neuron. (**B**) Ca^{2+} traces of the soma and five dendritic branches of the neuron shown in (A). Blue vertical lines indicate tone presentations, gray dots indicate detected Ca^{2+} transient, and arrowheads indicate compartment-specific Ca^{2+} transients. (**C**) Example of automatically detected "soma + dendrite" or "dendrite-only" Ca^{2+} transients. (**D**) Normalized Ca^{2+} transient amplitude correlation between the soma and dendrites in (A) (*r*, Pearson's correlation). (**E**) Soma-dendrite coupling decreases with distance from the soma. *r*, Pearson's correlation between somatic and dendritic Ca^{2+} traces (mean ± SEM, $n = 113$ soma-dendrite pairs from 13 neurons from seven mice). (**F**) Normalized amplitude in somas during Ca^{2+} transients detected in dendrites. For each distribution, events with low (<0.2) or high (>0.2) normalized amplitude were fitted with a Gaussian model. (**G**) Lower amplitude in somas during dendrite only transients ($P < 0.001$, Wilcoxon rank-sum test, $n = 215$ soma-dendrite pairs). (**H**) Normalized Ca^{2+} event amplitude in dendrites. Each distribution is fitted with a Gaussian model. (**I**) Lower amplitude in dendrites during dendrite only transients ($P < 0.001$, Wilcoxon rank-sum test, $n = 215$ soma-dendrite pairs). (**J**) Number of dendrites from a cell participating to soma + dendrite transients ($n = 113$ dendrites from 13 neurons from seven mice). (**K**) Number of dendrites from a cell participating to dendrite-only transients. (**L**) Number of dendrites from a cell participating to Ca^{2+} transients ($P < 0.001$, Wilcoxon rank-sum test). (**M**) Proportion of dendrites from a cell participating to transients co-occurring at the soma and in at least one dendrite. (**N**) Proportion of dendrites from a cell participating to dendrite-only transients. A small proportion of recorded dendrites from a cell are co-active during dendrite-only transients. (**O**) Proportion of dendrites from a cell participating to Ca^{2+} transients ($P < 0.001$, Wilcoxon rank-sum test). (**P**) Higher rate of dendrite-specific transients during tone presentation (soma + dendrite, $P = 0.020$ Wilcoxon signed-rank test; dendrite only, $P = 0.008$ Wilcoxon signed-rank test, $n = 113$ dendrites from 13 neurons from seven mice). (**Q**) Similar normalized amplitude of dendritic Ca^{2+} transients during spontaneous and tone-locked soma/dendrite transients ($P = 0.570$, Student's *t* test). (**R**) Similar normalized integral tone response of dendritic Ca^{2+} transients during spontaneous and tone-locked soma-dendrite transients ($P = 0.984$, Wilcoxon signed-rank test). (**S**) Higher amplitude of tone-locked than spontaneous dendrite-only transients ($P < 0.001$, Wilcoxon signed-rank test, $n = 53$ dendrites from 13 neurons). (**T**) Similar integral of tone-locked compared with spontaneous dendrite-only transients ($P = 0.071$, Student's *t* test).

To better interpret changes in sensory stimulus representation after learning, we first examined the stability of somatic CS responses across days in naïve animals under head-fixed two-photon imaging conditions. The somatic tone responsiveness was strongly conserved on up to 4 consecutive days, indicating a high cross-day stability of CS responses in LA PNs of nonconditioned animals (fig. S5).

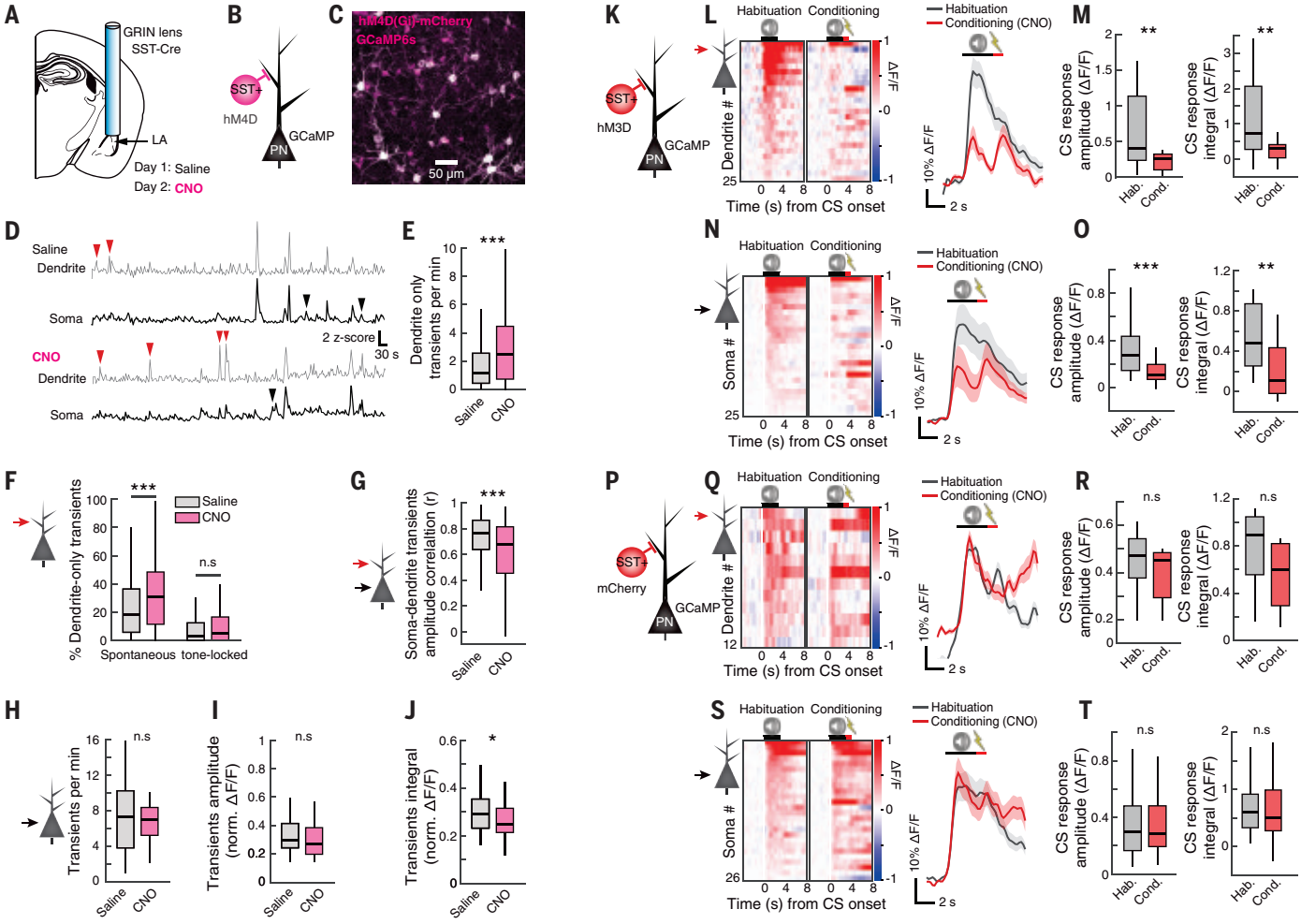


Fig. 3. SST+ interneurons control dendritic activity and CS responses. (**A**) Viral approach to express hM4D(Gi) in SST+ INs and GCaMP6s in a sparse population of LA PNs. Ninety tones were presented after saline or intraperitoneal injection of CNO on 2 consecutive days. (**B**) Experiment scheme. (**C**) Confocal image of hM4D(Gi)-mCherry in SST+ INs with sparse GCaMP6s in LA PNs. (**D**) Reduction of SST+ IN output increases dendrite-only transients. Arrowheads indicate compartment-specific transients in an example soma-dendrite pair. (**E**) Reduction of SST+ IN output increases the rate of dendrite-only transients ($P < 0.001$ Wilcoxon signed-rank test, $n = 79$ dendritic segments, 14 neurons). (**F**) Reduction of SST+ IN output increases spontaneous, but not tone-locked, dendrite-only transients (Spontaneous, $P < 0.001$; tone-locked, $P = 0.125$, Wilcoxon signed-rank test). (**G**) Reduction of SST+ IN output reduces the amplitude correlation between somatic and dendritic transients (r, Pearson's correlation coefficient, $P < 0.001$ Wilcoxon signed-rank test). (**H**) Reduction of SST+ IN output does not change the rate of somatic transients ($P = 0.347$, Wilcoxon signed-rank test, $n = 30$ neurons). (**I**) Reduction of SST+ IN output does not change the amplitude of somatic transients ($P = 0.140$, Student's t test). (**J**) Reduction of SST+ IN output decreases the integral of somatic transients ($P = 0.022$, Student's t test). (**K**) Experimental scheme. (**L**) (Left) CS responses ordered according to amplitude during habituation. (Right) Mean responses ± SEM ($n = 25$ CS-responsive dendrites during habituation, five mice). (**M**) Statistical analysis of (L). (Left) CS response $\Delta F/F$ amplitude ($P = 0.002$ Wilcoxon matched-pairs signed-rank test). (Right) CS response $\Delta F/F$ integral ($P = 0.705$ Wilcoxon matched-pairs signed-rank test; conditioning, 0.307 ± 0.096 $\Delta F/F$ integral). (**N**) (Left) CS responses ordered according to amplitude during habituation. (Right) Mean responses ± SEM ($n = 25$ CS-responsive somas during habituation, 5 mice). (**O**) Statistical analysis of (N). (Left) CS response $\Delta F/F$ amplitude ($P < 0.001$ Wilcoxon matched-pairs signed-rank test). (Right) CS response $\Delta F/F$ integral ($P = 0.002$ Wilcoxon matched-pairs signed-rank test). (**P**) Experimental scheme. (**Q**) (Left) CS responses ordered according to amplitude during habituation. (Right) Mean responses ± SEM ($n = 12$ CS-responsive dendrites during habituation, four mice). (**R**) Statistical analysis of (Q). (Left) CS response $\Delta F/F$ amplitude ($P = 0.979$ Student's t test). (Right) CS response $\Delta F/F$ integral ($P = 0.733$ Wilcoxon matched-pairs signed-rank test). (**S**) (Left) CS responses ordered according to amplitude during habituation. (Right) Mean responses ± SEM ($n = 26$ CS-responsive somas during habituation, four mice). (**T**) Statistical analysis of (S). (Left) CS response $\Delta F/F$ amplitude ($P = 0.585$ Student's t test). (Right) CS response ΔF/F integral ($P = 0.568$ Wilcoxon matched-pairs signed-rank test).

We then performed a 3-day differential auditory FC paradigm in a separate group of mice while recording somatic and dendritic Ca^{2+} activity in head-fixed mice under a two-photon microscope (Fig. 4A). Animals exhibited increased levels of freezing behavior when exposed to the CS 24 hours after FC under freely moving conditions before imaging Ca^{2+} responses (Fig. 4B). In these animals ($n = 9$), we imaged in total 177 neurons, 82% of which showed CS responses before or after conditioning, whereas 18% of all neurons were not CS responsive (fig. S6C). Compared with pre-conditioning levels, 84% of all CS responsive neurons (123 neurons from nine animals) exhibited learning-associated plasticity of CS responses, whereas 13% showed stable CS responses (fig. S6C).

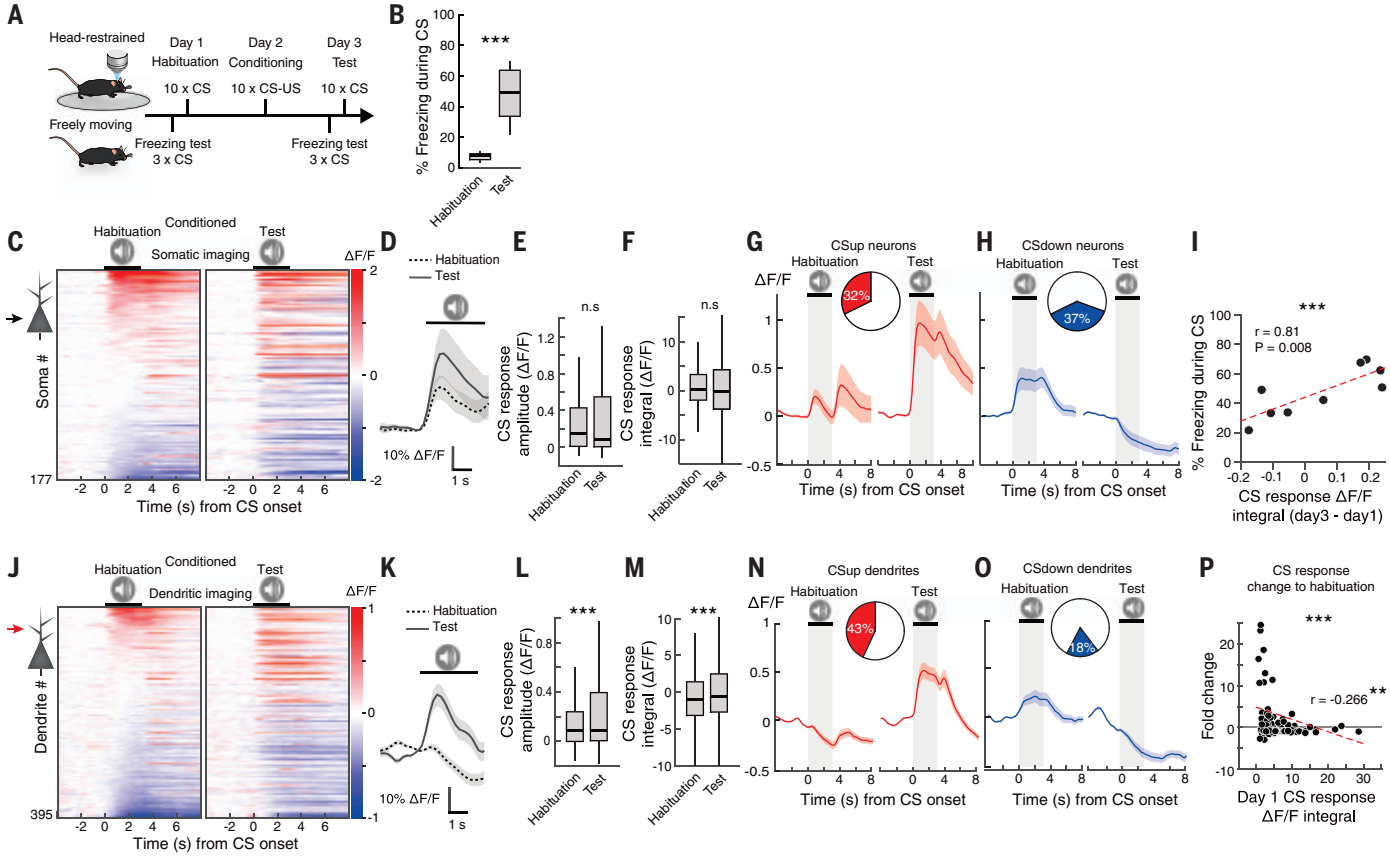


Fig. 4. CS response dynamics in LA somas and dendrites during auditory FC. (**A**) Auditory FC paradigm with simultaneous two-photon imaging. Freezing tests were performed in freely moving conditions. (**B**) Mice learn the CS-US association. Shown is percent time spent freezing, during the CS before and after learning ($P < 0.001$, Wilcoxon signed-rank test; habituation, $n = 6$ mice; test, $n = 9$ mice). (**C**) CS responses in somas before and after FC ordered according to amplitude during habituation ($n = 177$ somas from nine mice). (**D**) Mean somatic CS response before and after FC ($n = 177$ somas from nine mice, mean ± SEM). (**E**) The mean somatic CS response amplitude is similar before and after FC ($P = 0.641$ Wilcoxon signed-rank test). (**F**) The mean somatic CS response integral is similar before and after FC ($P = 0.547$ Wilcoxon signed-rank test). (**G** and **H**) A large proportion of somas (G) up-regulate and (H) down-regulate their CS response upon FC (mean ± SEM). (**I**) Time spent freezing during CS is correlated with somatic CS response integral change after FC (r, Pearson's correlation, $P = 0.008$, $n = 177$ somas, 19.6 ± 5.4 neurons per mouse from nine mice). (**J**) CS responses in dendrites before and after FC ordered according to amplitude during habituation ($n = 395$ dendrites from nine mice). (**K**) Mean CS response of all dendrites before and after FC (mean ± SEM). (**L**) The mean dendritic CS response amplitude increases after FC ($P < 0.001$ Wilcoxon signed-rank test). (**M**) The mean dendritic CS response integral increases after FC ($P < 0.001$ Wilcoxon signed-rank test, $n = 395$ dendrites from nine mice). (**N**) A large proportion of dendrites up-regulate and (**O**) a smaller proportion down-regulate their CS response after FC (mean ± SEM). (**P**) Dendritic CS response fold change after paired conditioning normalized to habituation. The median response integral is similar ($P = 0.172$ one-sample Wilcoxon signed-rank test, 173 ± 57% CS response integral increase). The CS response integral is anticorrelated to the CS response integral during habituation (r, Pearson's correlation coefficient, $n = 97$ dendrites from nine mice).

Whereas the overall average somatic CS-response magnitude remained similar after FC (Fig. 4, C to F), neurons could be subdivided into three functional subpopulations: neurons exhibiting increased CS responses (CSup neurons; 32% of $n = 57$ of all neurons from $n = 9$ animals), neurons exhibiting decreased CS responses (CSdown neurons; 37% of all neurons), and neurons with stable CS responses (Fig. 4, G and H).

Similar proportions of CSup and CSdown neurons were US responsive or US nonresponsive ($P = 0.527$, Fisher's exact test) (fig. S6, F and G), indicating that somatic US responsiveness is not necessary for CS response plasticity, nor does it predict the direction of plasticity.

To test the specificity of the observed FC-induced plasticity of CS responses, we subjected mice to an unpaired conditioning paradigm, in which the CS and the US were presented independently (more than 120 s apart) (fig. S7). Unpaired conditioned mice did not freeze when exposed to the CS (fig. S7B). Consequently, unpaired conditioning resulted in an overall reduction of somatic CS response amplitudes at the population level (fig. S5, C to F, and fig. S9).

Comparing the proportions of CSup and CSdown neurons between naïve mice and mice subjected to associative conditioning or unpaired conditioning protocols revealed that CSup neurons were overrepresented in conditioned mice relative to mice subjected to unpaired conditioning ($P < 0.001$, Fisher's exact test), whereas CSdown neurons were found in similar proportions in animals that underwent paired or unpaired conditioning ($P = 0.247$, Fisher's exact test) (Fig. 4, G and H, and fig. S7, G and H).

The absolute CS response change during auditory FC strongly correlated with acquired freezing behavior after conditioning in animals subjected to paired (Fig. 4I), but not unpaired, conditioning (fig. S7I), which suggests that updating of CS response representations is necessary for conditioned freezing behavior (*21*). This correlation relied on both CSup and CSdown neuron activity (fig. S7, Q and R), which suggests that both CS response

plasticity types contribute to the new CS representation in the amygdala after conditioning.

Associative FC induces differential somatic and dendritic plasticity

We next examined FC-induced dendritic plasticity. In total, we imaged 395 dendrites (from nine animals), 75% of which showed CS responses before or after conditioning, whereas 25% of all dendrites were not CS responsive. Compared with preconditioning levels, 87% of all CS responsive dendritic segments (258 dendritic segments from nine animals) exhibited learning-associated plasticity of CS responses, whereas 13% showed stable CS responses (fig. S6C).

Overall, FC induced both up- and down-regulation of CS responses in dendrites (Fig. 4, J, N, and O). However, in contrast to somatic plasticity, FC resulted in an overall net increase in dendritic CS responses (Fig. 4, K to M) driven by a large fraction of CSup dendrites (43%), with only a relatively small fraction of CSdown dendrite (18%) (Fig. 4, N to P). There was a higher proportion of CSup than CSdown dendritic segments, in contrast to the almost equal proportions observed at the somatic level (fig. S6C).

To assess whether the overall potentiation of dendritic CS responses was specific for associative FC, we analyzed dendritic CS response plasticity in mice subjected to an unpaired conditioning paradigm. The dendritic CS response was decreased in unpaired conditioned animals (fig. S7, K to M). These results were robust to normalization to the mean initial CS response amplitude between groups (fig. S10). Similar to the changes of somatic CS responses upon unpaired conditioning, there was a large fraction of CSdown dendrites (35%), with only a minor population of CSup dendrites (7%) (fig. S7, N to P).

CS response plasticity in dendritic spines

Consistent with the large fraction of CSup dendrites, CS responses of dendritic spines were on average increased (fig. S8, A to E). Similar to somatic and dendritic responses, spines showed CSup and CSdown plasticity patterns (fig. S8, F to I). In total, we recorded 112 spines on identified CSup dendrites and 28 spines on identified CSdown dendrites. On average, CSup dendrites exhibited more spines with increased CS responses (CSup spines) than spines with decreased CS responses (CSdown spines) (fig. S8J). By contrast, CSup spines were underrepresented on CSdown dendrites (fig. S8K).

Consistent with the idea of anatomical clustering of dendritic spines with a conditioning-induced increase in their CS responses (CSup spines), the distance between CSup spines located on a given dendritic branch was found to be smaller as compared with the distance between CSdown spines (fig. S8L). Last, our results indicate that US-responsive dendrites contain a comparatively larger fraction of CSup spines, whereas US-unresponsive dendrites contain more CSdown spines (fig. S8M).

Subcellular compartment-specific plasticity in single neurons

Given the variability of soma-dendrite correlation in single neurons (Fig. 2, B and D), we next examined whether learning would change the variability of CS responses between individual dendritic branches belonging to the same neuron. CS response plasticity was variable between the different dendritic branches of single neurons (Fig. 5, A, B, and G), and somatic and dendritic activity became more decorrelated after FC (Fig. 5, C and D). The amplitude decorrelation of somatic and dendritic transients was more pronounced in dendrites located several branch points away from the soma—third order or higher as compared with first- and second-order branches (fig. S9B). However, we did not observe a linear effect of distance from soma on transient amplitude correlation (fig. S9A).

The overall probability of a CS input to trigger local, dendrite-only transients was increased upon learning (Fig. 5E), but the number and proportion of dendrites co-active during somatic activity did not change upon conditioning (fig. S9, C to P), indicating that plasticity of CS responses can be restricted to individual dendritic branches. In contrast to CS-evoked responses, the proportion of dendrite-only events during spontaneous activity did not change upon FC (fig. S9, Q to S).

Dendritic CS response variability increased in conditioned mice but was not observed upon unpaired conditioning (Fig. 5G), suggesting that associative learning specifically induces decorrelation of somatic and dendritic CS responses.

Most dendrites of somatically US-responsive neurons also exhibited US responses (44 of 79 dendrites, 56%). Localized dendritic US responses were also present in some dendrites of neurons that lacked somatic US responses (35 of 158 dendrites, 22%) (fig. S6, H and I), suggesting that dendrite-specific US responses might still influence CS response plasticity at the soma.

To examine how dendritic CS response plasticity relates to somatic plasticity in single neurons, we compared the plasticity of individual dendritic branches with somatic plasticity in identified CSup and CSdown neurons (Fig. 5, H to S). The amplitude of dendritic CS responses increased in CSup neurons (Fig. 5, N to P). By contrast, dendritic and somatic CS response plasticity was uncoupled in CSdown neurons. Whereas the mean amplitude of somatic CS responses was decreased in CSdown neurons, CS responses increased in the dendrites of the same neurons (Fig. 5, Q to S), indicating that the cellular mechanisms underlying CS response plasticity in CSup and in CSdown neurons are fundamentally different.

Perisomatic inhibition uncouples dendritic and somatic CS response plasticity

One possible mechanism that might underlie the uncoupling of somatic and dendritic CS response plasticity observed in CSdown neurons could be a conditioning-induced increase in perisomatic inhibition (*26*). In the LA, like in the cortex or hippocampus, parvalbumin (PV)–expressing interneurons mediate perisomatic inhibition and tightly control PN activity (*21*). To investigate the effect of PV+ interneurons on somatic CS response plasticity, we chemogenetically suppressed their activity (Fig. 6, A to C). Inhibition of LA PV+ interneurons led to a reduced inhibitory drive onto PNs in vitro (Fig. 6, D to F) and to an increased rate of somatic Ca^{2+} transients in PNs in vivo (Fig. 6G). Inhibition of PV+ interneurons during memory retrieval (test session) 24 hours after FC decreased the proportion of CSdown neurons compared with that in control animals, without any effect on the proportion of CSup neurons (Fig. 6J). In the remaining CSdown neurons that we detected while inhibiting PV+ interneurons, the CS response amplitude decreased, however, to a lesser extent than in the control group, whereas CS response amplitudes increased similarly in CSup neurons (Fig. 6, K to N).

Discussion

We used in vivo chronic two-photon Ca^{2+} imaging to investigate dendritic function and plasticity in LA PNs. Under baseline conditions, somatic and dendritic Ca^{2+} transients were strongly coupled. Recent studies in the neocortex reported that the fraction of Ca^{2+} transients restricted to dendrites is relatively small (*10*–*12*). Pervasive global Ca^{2+} transients could reflect efficient somatic AP back-propagation into the dendritic arbor of LA PNs (*10*, *27*). However, the temporal resolution of the currently available Ca^{2+} sensors does not permit to determine the temporal sequence and causality of somatic and dendritic activity. Nonetheless, in the LA the correlation between the amplitude of co-occurring somatic and dendritic Ca^{2+} transients was variable and decreased with increasing distance from the soma within single neurons, and local dendritic activity occurred independently from somatic activity (Fig. 2). The rate and amplitude of dendrite-specific Ca^{2+} transients was notably higher during auditory sensory stimulation (Fig. 2, P and S), indicating that dendrites of LA PNs locally integrate auditory inputs.

In cortex and in the basolateral amygdala (BLA), dendritic activity is thought to be regulated by dendrite-targeting interneurons

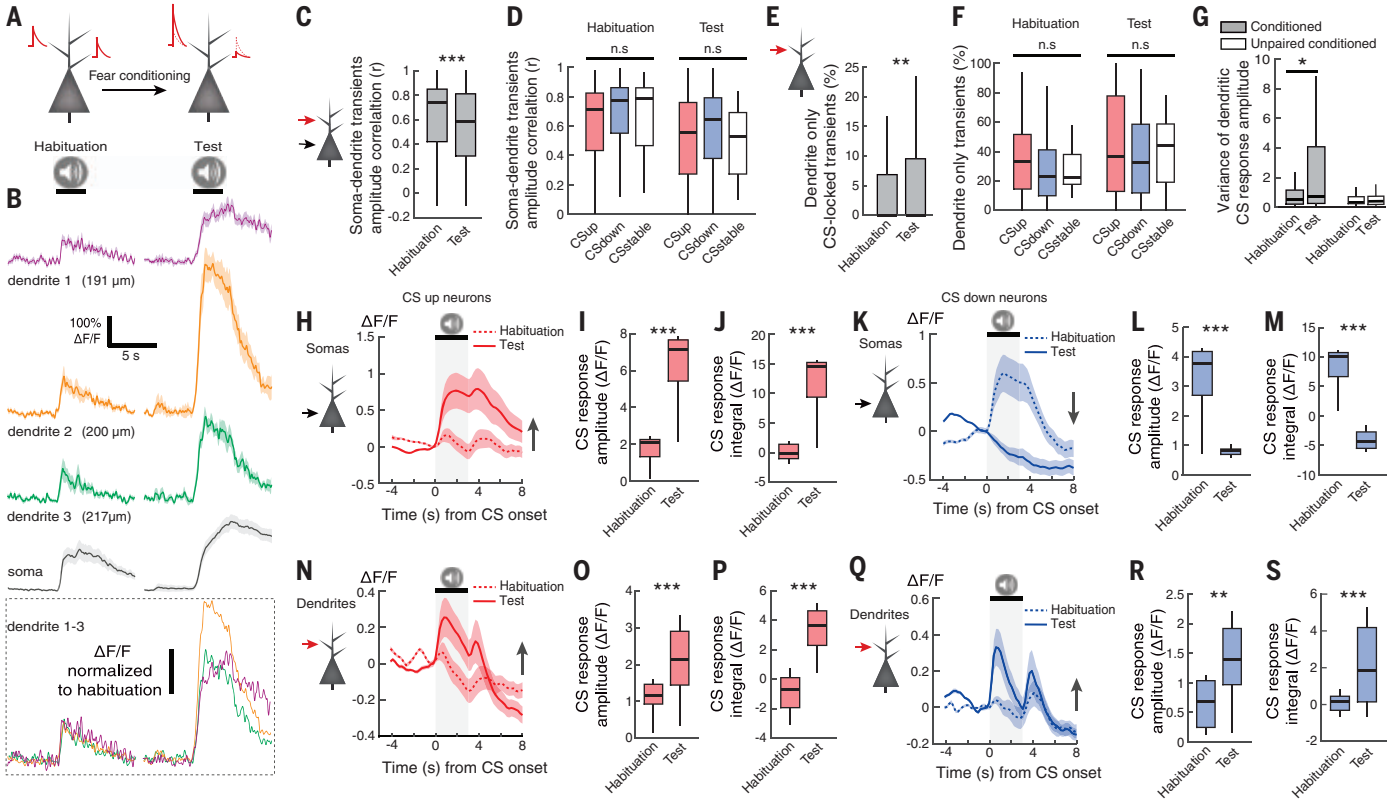


Fig. 5. Uncoupling of somatic and dendritic plasticity of CS response upon auditory FC. (**A**) Simultaneous imaging of the soma and dendrites. (**B**) Variability of CS response between the soma and three dendritic segments of a neuron before and after FC (mean ± SEM). (Bottom) CS responses normalized to habituation day. (**C**) Correlation of soma-dendrite event amplitude decreases after FC (*r*, Pearson's correlation coefficient, $P < 0.001$ Wilcoxon signed-rank test, $n = 262$ soma-dendrite pairs, 34 neurons). (**D**) Correlation of soma-dendrite transient amplitudes is similar between CSup, CSdown, and CSstable neurons [*r*, Pearson's correlation coefficient, habituation: $F(2, 185) = 2.34$, $P = 0.099$; test, $F(2, 185) = 1.74$, $P = 0.179$; both one-way ANOVA; $n = 69$ soma-dendrite pairs, 15 CSup neurons, 199 soma-dendrite pairs, 21 CSdown neurons, 19 soma-dendrite pairs, seven CSstable neurons]. (**E**) The proportion of CS-locked dendrite-only transients increases after FC ($P = 0.008$ Wilcoxon signed-rank test, $n = 262$ soma-dendrite pairs, 34 neurons). (**F**) The proportion of CS-locked dendrite-only transients is similar between CSup, CSdown, and CSstable neurons [habituation, $F(2, 185) = 2.34$, $P = 0.099$; test, $F(2, 185) = 1.74$, $P = 0.179$; both one-way ANOVA]. (**G**) The variance of CS response amplitude increases between dendrites after FC ($P = 0.010$, both Student's *t* test, $n = 250$ dendrites, 50 neurons, nine mice). (**H**) Mean CS response increases in CSup somas after FC ($n = 19$ neurons, mean ± SEM). (**I**) CS response amplitude increases in CSup somas after FC ($P < 0.001$, Wilcoxon signed-rank test). (**J**) CS response integral increases in CSup somas after FC ($P < 0.001$, Student's *t* test). (**K**) Mean CS response decreases in CSdown somas after FC ($n = 18$ neurons, mean ± SEM). (**L**) CS response amplitude decreases in CSdown somas after FC ($P < 0.001$, Wilcoxon signed-rank test). (**M**) CS response integral decreases in CSdown somas after FC ($P < 0.001$, Wilcoxon signed-rank test). (**N**) Mean CS response increases in the dendrites of CSup neurons after FC ($n = 110$ dendrites, mean ± SEM). (**O**) CS response amplitude increases in the dendrites of CSup neurons after FC ($P < 0.001$, Wilcoxon signed-rank test). (**P**) CS response integral increases in the dendrites of CSup neurons after FC ($P < 0.001$, Student's *t* test). (**Q**) Mean CS response increases in the dendrites of CSdown neurons after FC ($n = 70$ dendrites, mean ± SEM). (**R**) CS response amplitude increases in the dendrites of CSdown neurons after FC ($P = 0.004$ Wilcoxon signed-rank test). (**S**) CS response integral increases in the dendrites of CSdown neurons after FC ($P = 0.002$, Student's *t* test).

(*28*–*30*). SST+ interneurons, in particular, predominantly contact PN dendrites and provide input-specific inhibition to PNs. SST+ interneurons in the BLA are inhibited during CS presentation, putting them in an ideal position to control dendritic excitability and gate dendritic integration and plasticity (*23*, *24*). CS-induced inhibition of SST+ interneurons is necessary for learning (*24*). We found that dendrite-specific Ca^{2+} transients are under the control of SST+ interneurons in vivo and that silencing of SST+ interneurons increases dendrite-specific transients without affecting somatic activity in PNs (Fig. 3, E to J). Moreover, SST+ interneurons suppress spontaneous, less salient inputs to PN dendrites (Fig. 3F). The relief of SST+ interneuron–mediated dendritic inhibition is necessary for the generation of dendritic CS responses during conditioning (Fig. 3, L and M). This is consistent with the notion that disinhibition through SST+ interneurons creates a temporal window that allows for the induction of dendritic plasticity during the coincident US (*23*, *24*).

Upon FC, the somatic CS response of LA PNs remained stable on the population level (Fig. 4, D to F). However, single-neuron dynamics revealed that bidirectional plasticity of somatic CS responses (Fig. 4, G and H) and freezing behavior was correlated with the absolute somatic CS response change (Fig. 4I), suggesting that CSup and CSdown neurons contribute not only to CS discrimination (*21*) but also to conditioned freezing behavior after FC. Our observation that the correlation between CS response plasticity and conditioned freezing behavior depended on both CSup as well as CSdown neurons may suggest that these two populations might function in concert to induce appropriate behavioral responses.

By contrast, LA PN dendrites and spines predominantly up-regulated their CS response upon FC (Fig. 4, K to M and P, and fig. S8, A to E), indicating, on average, a higher proportion of potentiated CS responses in dendrites. These changes were specifically induced in animals presented with temporally contingent

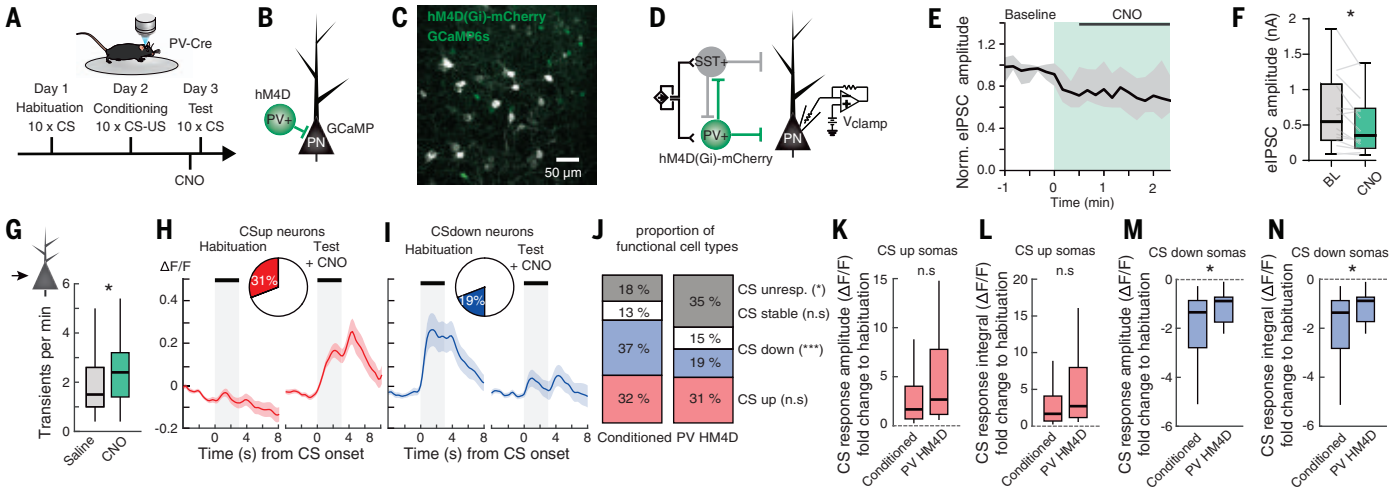


Fig. 6. Uncoupling of dendritic and somatic plasticity in CSdown neurons is mediated by an increase in perisomatic inhibition. (**A**) Auditory FC paradigm with simultaneous two-photon imaging. (**B**) Experiment scheme. (**C**) Confocal microscopy image of hM4D(Gi)-mCherry in PV+ INs with sparse GCaMP6s in LA PNs. (**D**) Whole-cell recording of evoked inhibitory postsynaptic current (eIPSC) in BLA PNs. (**E**) Reduction of PV+ IN output effect on eIPSC amplitude in BLA PNs (n = 14 cells, three mice). Median (black) and interquartile range (gray) are normalized to median baseline period. Green shaded area indicates bath application of CNO. Top black line indicates the CNO analysis window. (**F**) Statistical analysis of (E) (P = 0.017, Wilcoxon matched-pairs signed-rank test). (**G**) Suppression of PV+ IN output increases the rate of somatic transients (P = 0.012, one-sample Wilcoxon signed-rank test, n = 66 neurons). (**H**) A large proportion of somas up-regulate and others (**I**) down-regulate their CS response after FC (mean ± SEM). (**J**) Similar proportion of CSup and smaller proportion of CSdown somas in mice with suppressed PV+ IN activity in the amygdala (CSup, P = 0.618; CSdown, P < 0.001; both Fisher's exact test, Bonferroni corrected; conditioned, n = 177 neurons from nine mice; PV CNO, n = 66 neurons from six mice). (**K**) Reduction of PV+ IN output does not affect CS response amplitude change in CSup neurons during auditory FC (P = 0.552 Wilcoxon rank-sum test; conditioned, n = 57 neurons; PV CNO, n = 39 neurons). (**L**) Reduction of PV+ IN output does not affect CS response integral change in CSup neurons during auditory FC (P = 0.622 Wilcoxon rank-sum test; conditioned, n = 57 neurons; PV CNO, n = 39 neurons). (**M**) Reduction of PV+ IN output decreases the CS response amplitude change of CSdown neurons after auditory FC (P = 0.039 Wilcoxon rank-sum test; conditioned, n = 66 neurons; PV CNO, n = 27 neurons). (**N**) Reduction of PV+ IN output decreases the CS response integral change of CSdown neurons after auditory FC (P = 0.032, Wilcoxon rank-sum test; conditioned, n = 66 neurons; PV CNO, n = 27 neurons).

CS-US pairings. By contrast, in unpaired conditioned mice exposed to equal CSs and USs, but in a temporally unpaired manner, CS responses of somas and dendrites decreased uniformly (fig. S7, C to F and J to M). CS-US pairing thus induces dendritic plasticity of CS responses that is not fully reflected at the level of the soma. To explore the relationship between somatic and dendritic CS response plasticity at the single-neuron level, we simultaneously examined conditioning-induced changes in somas and in connected dendrites. Consistent with the population-level analysis and with previous one-photon imaging results from populations of amygdala neurons in freely moving animals (*20*, *21*), we found that LA PN somas could be divided into three functional plasticity types with either increased, decreased, or unchanged CS responses. Only a minority of CSup and CSdown somas exhibited converging CS and US responses during conditioning, indicating that both the up-regulation and the down-regulation of somatic responses may result from different induction processes involving synaptic and/or dendritic plasticity as well as network effects.

Our finding that individual dendritic branches can exhibit US responses that do not result in somatic transients and are uncoupled from somatic activity suggests that conditioning may lead to local, compartmentalized plasticity in dendrites that might encode the CS-US association memory in the BLA (*27*). CSup and CSdown spines are preferentially located on CSup and CSdown dendrites, respectively (fig. S8). Further, consistent with the idea of anatomical clustering of potentiated synaptic inputs (*31*), the distance between CSup spines located on a given dendritic branch is smaller as compared with the distance between CSdown spines (fig. S6L). Therefore, conclusions based on somatic US responses may underestimate the complexity of dendritic integration and plasticity as a source of neuronal output adaptation and learning (*32*), as suggested by a recent computational modeling study (*33*).

We observed that FC resulted in a decorrelation of CS responses between dendrites and their parent soma, as well as between individual dendritic branches belonging to the same neuron (Fig. 5, B, C, and E). This may reflect conditioning-induced and branch-specific changes in action potential backpropagation, possibly resulting from changes in SST+ interneurons–mediated dendritic inhibition. More likely, FC might lead to the induction of localized synaptic and dendritic CS response plasticity. The more pronounced decorrelation observed in more distal, higher-order dendrites (Fig. 5C) argues for local dendritic plasticity as the origin of soma-dendrite decorrelation than for alterations in action potential backpropagation that would increase soma-dendrite correlations. Moreover, although the proportion of dendrite-only CS responses increased after FC (Fig. 5E), the overall proportion of spontaneous dendrite-only Ca^{2+} transients did not change (fig. S9, Q to S). This suggests that FC does not lead to a general increase in dendritic excitability but specifically increases dendritic responses to the conditioned sensory inputs (*34*).

Whereas dendrites and their parent soma of CSup neurons generally exhibited a correlated increase in their CS responses after learning, this was not the case for CSdown neurons. Even though the somas of CSdown neurons showed decreased CS responses upon learning, dendrites belonging to the same neurons did not but instead exhibited increased CS responses. This strongly suggests that in CSdown neurons, learning-induced dendritic and somatic plasticity mechanisms are distinct from CSup neurons.

Although increased dendritic responses in CSdown neurons likely result from potentiation of synaptic inputs, decreased somatic responses may reflect soma-specific changes of neuronal excitability (*35*–*37*) and/or plasticity of local soma-targeting inhibitory circuits (*28*, *38*).

In support of the latter, suppressing the activity of soma-targeting PV+ interneurons reduced the proportion of CSdown neurons after FC (Fig. 6J), thus suggesting that in CSdown neurons, increased perisomatic inhibition counteracts excitatory dendritic inputs (*39*). Consistent with this, somatic inhibition undergoes anatomical and functional plasticity upon FC, resulting in a network-wide remodeling of the excitation/inhibition balance (*26*, *40*).

Stimulus-specific dendritic response strengthening along with changes in stimulus-specific somatic inhibition could be a hallmark of network reorganization during the formation of new associations that allow stimulus-specific behavioral responses or the regulation of stimulus discrimination after learning or contribute to circuit homeostasis (*21*, *41*, *42*). Classical Hebbian learning models neither consider the contribution of inhibitory interneurons and the subcellular compartments they target nor the differential plasticity of distinct neuronal compartments and therefore fail to appropriately describe effects on the level of local microcircuits (*43*, *44*). It is still unclear which factors determine the direction of plasticity in single neurons. Studies point to the importance of a neuron's intrinsic excitability at the time of memory formation (*45*), its afferent inputs and/or output targets (*18*, *46*, *47*), the role of inhibitory interneurons (*23*, *24*, *38*), and the clustering of synaptic inputs along the dendritic arbor (*38*, *48*). Future experiments studying how specific long-range inputs and inhibitory microcircuits control somatodendritic coupling as well as compartment-specific signal integration and plasticity might reveal general principles of network reorganization underlying the encoding and storage of newly formed memories.

Materials and methods

Animals

All animal procedures were performed in accordance with institutional guidelines at the Friedrich Miescher Institute for Biomedical Research and were approved by the Veterinary Department of the Canton of Basel-Stadt. SOM-ires-Cre (*49*) and PV-ires-Cre mice (*50*) were used for Cre-dependent expression of viral vectors. Only heterozygous (Cre/wild type) mice were used for experiments and were backcrossed to a C57BL/6J background for more than ten generations. For all other experiments, male wild-type C57BL/6J mice (Envigo, the Netherlands) were used. Mice were individually housed for at least 14 days before starting behavioral paradigms. Animals were kept in a 12 hours light–dark cycle with access to food and water ad libitum as well as a shelter and a running wheel. All behavioral experiments were conducted during the light cycle.

Surgical procedure

We performed surgeries when mice were 9 to 10 weeks of age. Mice were anesthetized using isoflurane (3 to 5% for induction, 1 to 2% for maintenance; Attane, Provet) in 95% O_2 (Praxair) and fixed in a stereotactic frame (Kopf Instruments). Injections of buprenorphine (Temgesic, Indivior UK Limited; 0.1 mg per kg body weight subcutaneously 30 min before anesthesia) and ropivacain (Naropin, AstraZeneca; 0.1 ml locally under the scalp before incision) were provided for analgesia. Postoperative pain medication included buprenorphine (0.1 mg per kg body weight in the drinking water; overnight) and injections of meloxicam (Metacam, Boehringer Ingelheim; 1 mg per kg body weight subcutaneously) for up to 3 days if necessary. Ophthalmic ointment was applied to avoid eye drying. The body temperature of the experimental animal was maintained at 36°C using a feedback-controlled heating pad (FHC).

For the imaging of somas and dendrites, mice were injected with AAV2/9-CaMKII-Cre-SV40 ($2.80.10^{13}$ GC.ml^{-1}, AV-9-PV2396, UPenn) diluted in sterile saline (final dilution 1:20000) mixed 1:1 with AAV2/1-Syn-Flex-GCaMP6s (*51*) ($1.9.10^{12}$ GC.ml^{-1}, 100845-AAV2/1, Addgene, final dilution: 1:2), resulting in sparse labeling of LA PNs. For combined imaging of LA pyramidal cells and chemogenetic inhibition of SST+ interneurons, SST-ires-Cre mice were injected with an AAV2/1-CaMKII-FLPo diluted in sterile saline ($3.4.10^{12}$ GC.ml^{-1}, Vector Biolabs, VB1921, final dilution 1:5000) mixed 1:1:1 with AAV2/1-Ef1a-fDIO-GCaMP6s ($4.2.10^{12}$ GC.ml^{-1}, 105714, Addgene, final dilution 1:3) and AAV2/1-hSyn-DIO-hM4D(Gi)-mCherry ($6.4.10^{12}$ GC.ml^{-1}, 44362-AAV2/2, Addgene, final dilution 1:3). For combined imaging of LA pyramidal cells and chemogenetic excitation of SST+ interneurons, SST-ires-Cre mice were injected with an AAV2/1-CaMKII-FLPo diluted in sterile saline ($3.4.10^{12}$ GC.ml^{-1}, Vector Biolabs, VB1921, final dilution 1:900) mixed 1:1:1 with AAV2/1-Ef1a-fDIO-GCaMP6s ($4.2.10^{12}$ GC.ml^{-1}, 105714, Addgene, final dilution 1:3) and AAV2/2-hSyn-DIO-hM3D(Gq)-mCherry ($7.9.10^{12}$ GC.ml^{-1}, 50474-AAV2/2, Addgene, final dilution 1:3). In the case of control mice, the h3MD(Gq)-mCherry expressing virus was replaced by an equivalent amount of AAV2/2-hSyn-DIO-mCherry ($3.9.10^{12}$ GC.ml^{-1}, 50459-AAV2/2, Addgene, final dilution 1:3). The virus mix (300-500 nl) was unilaterally injected into the BLA using a precision micropositioner (Model 2650, Kopf Instruments) and pulled glass pipettes connected to a Picospritzer III microinjection system (Parker Hannifin Corporation) at the following coordinates from bregma: AP: −1.55 mm; ML: −3.4 mm; DV: −3.75 to −4.15 mm. During the same surgery, a GRIN microendoscope (GRIN lens, 0.6 × 7.3 mm, GLP-0673, Inscopix) was implanted into the LA as previously described (*52*). In brief, a sterile needle was used to make an incision above the imaging site. The GRIN lens was subsequently lowered into the brain using a micropositioner (coordinates from bregma: AP: −1.55 mm; ML: −3.4 mm; DV: 4.0 mm) with a custom-built lens holder and fixed to the skull using ultraviolet light-curable glue (Henkel, Loctite 4305). A mix of dental acrylic (Paladur, Heraeus) and black acrylic paint was used to seal the skull and attach a custom-made head bar for animal fixation during imaging experiments. Mice were allowed to recover for at least 14 days after GRIN lens implantation before checking for GCaMP expression.

Two-photon imaging

Two to six weeks after surgery, mice were head-fixed on a running wheel to check for sufficient expression of GCaMP6s under a two-photon microscope (Ultima Investigator, Bruker, USA). When sufficient expression level was reached (5 to 10 weeks after surgery), mice were habituated to a brief head-fixation under the microscope for at least 2 consecutive days before any behavioral paradigm, while free to run on a wheel. GCaMP6s signal was recorded at a 30 Hz from a field of view of 471 by 471 μm (512 by 512 pixels) through a water immersion objective [25×, 1.05 numerical aperture (NA), Olympus]. Ultrasound gel (G008, FIAB spA) was used to interface the objective and the GRIN lens. Excitation light was provided with a mode-locked laser system operating at 920 nm, 80-MHz pulse repeat, 120 fs pulse width (Insight X3, Spectra Physics, Mountain View, CA). For simultaneous imaging of GCaMP6s and tdTomato (AAV-CAG-flex-tdTomato, $1.3.10^{13}$ GC.ml^{-1}, 28306-AAV1, Addgene), excitation light was provided at 960 nm instead and green and red emission light was filtered through specific bandpass filters (green emission light: 525/70 nm; red emission light: 595/50 nm, both Chroma Technology) and detected by different PMTs (photomultipliers, Hamamatsu Photonics). For individual mice, the same imaging parameters were kept across repeated behavioral sessions. Imaging data were recorded using Prairie View Software (Bruker, USA).

For detailed structural scan acquisition and subsequent dendritic arbor reconstruction, mice were hosted for an additional 4 to 6 weeks following any behavioral paradigm to allow for maximal GCaMP6s expression. Then, mice were anesthetized with FMM (Fentanyl-Curamed, 0.05 mg per kg of body weight, Midazolam, 5 mg per kg of body weight, Medetomidine, 0.5 mg per kg of body weight) injected IP (*53*). Ophthalmic ointment was applied to avoid eye drying. The body temperature of the experimental animal was maintained at 36°C using a feedback-controlled heating

pad (FHC). A 3D imaging volume, centered on the imaging plane previously used for functional imaging, was acquired with a 2 μm increment between each plane. Each image was obtained by averaging 128 frames from a field of view of 471 by 471 μm (1024 by 1024 pixels). The dendritic arbors of the imaged neurons were tracked in 3D using the ImageJ simple neurite tracer plugin (*54*), projected along the depth dimension and mapped onto the corresponding functional imaging data (fig. S2A). This allowed associating functionally imaged dendritic branches with their parent soma.

Mouse behavior

To assess tone and shock tuning and stability, mice were head-fixed on a running wheel under a 2-photon microscope. A rotary encoder (SparkFun Electronics) was used to record running speed. Mice were allowed to habituate to the setup under head fixation for at least 2 days prior to the start of any experimental paradigm and for at least 15 min prior to the beginning of each experimental session. For auditory stimulus delivery, an electrostatic speaker (ES1, Tucker-Davis Technologies) was placed on each side of the mouse's head. 3 s continuous pure tones (75 to 85 dB, 3 to 18 kHz) were generated using a System 3 RP2.1 real-time processor and a SA1 stereo amplifier with RPvdsEx Software (all Tucker-Davis Technologies). Auditory stimuli were presented in a random order of frequencies and with increasing sound pressure levels on 2 consecutive days with an inter-trial interval (ITI) of 90 s. In a subset of mice, the same tone presentation protocol was applied for 4 consecutive days. For electrical shock delivery, a small two pin connector (Fischer Elektronik) was brought in contact to the forehead of the mouse, between the eyes and the ears, and connected to a precision animal shocker delivering direct current (DC). On the last day, 20 to 30 min after the last tone presentation, 1 s electrical shocks with increasing intensities (0.1 to 0.65 mA) were delivered with an inter-trial interval (ITI) of 120 s. Behavioral protocols for stimulus control were generated using the Prairie View Software (Bruker, USA) via TTL (transistor-to-transistor logic) pulses. For chemogenetic reduction of SST+ interneurons synaptic strength, mice were injected IP with saline (0.9% NaCl) on day 1 and with CNO (5 mg per kg, Tocris) on day 2. 1 hour after CNO injection, tones were presented on both days following the same presentation protocol as described in this section.

FC paradigm

Mice were extensively handled and habituated to the head-fixation and electrodes delivering shock for 2-5 days prior to imaging or behavioral recording. On day 1, mice were habituated to a pure tone (CS: 3 s, 7 or 12 kHz; 80 to 85 dB sound pressure level) presented 10 times with a 90 to 120 s inter-trial interval (ITI). On day 2, mice were conditioned to the CS by pairing it with a US (1 s electrical shock, 1 mA) delivered as described in the above section. Alternatively, the US was delivered with two small electrodes brought in contact to the skin of the left forelimb and left hindlimb of the mouse and connected to a precision animal shocker delivering direct current (1 s electrical shock, 1 mA, DC). The US was applied directly after the CS. CSs and USs were each presented 10 times with a 120 s inter-trial interval (ITI). In a separate group of mice that underwent unpaired conditioning, the US was applied at least 120 s apart from any CS. On day 3, similar to day 1, the same pure tones were presented 10 times with a 90 to 120 s inter-trial interval (ITI). For FC paradigm combined with chemogenetic manipulation of interneurons activity, SOM-ires-Cre mice were injected IP with CNO or saline on day 2 and PV-ires-Cre mice on day 3, 30 to 90 min before two-photon imaging.

Freely moving behavior

In parallel with head-fixed FC under a two-photon microscope, fear memory was tested in freely moving conditions in the same animals. On day 1 and 3, mice were placed in a circular arena with Plexiglas walls in a sound-isolated box. After a 2 min baseline period, 3 CSs (30 s, continuous pure tone, 7 or 12 kHz, 80-85 dB sound pressure level) were presented with a 60 to 120 s pseudorandom inter-trial interval (ITI). Freezing responses were quantified on day 1 (habituation) and day 3 (test) during CS presentations. Behavioral protocols for stimulus control were generated using the Radiant software (Plexon) or RPvdsEx software (Tucker Davis Technologies) via TTL (transistor-to-transistor logic) pulses. The animal motion was tracked and freezing during the CS presentations on day 1 and day 3 was compared to evaluate associative learning. Due to technical issues during recording, freezing was not measured during day 1 habituation in 3 fear conditioned mice and during day 3 test in one mouse in the unpaired conditioned group. N numbers are indicated in figure legends.

In vitro electrophysiology

Acute brain slice preparation

Mice were deeply anesthetized (ketamine 250 mg per kg and medetomidine 2.5 mg per kg body weight intraperitoneally) and perfused with carbogenated (95% O_2, 5% CO_2) ice-cold slicing solution ([mM] 2.5 KCl, 11 glucose, 234 sucrose, 26 $NaHCO_3$, 1.25 NaH_2PO_4, 10 $MgSO_4$, 2 $CaCl_2$; 340 mOsm). After decapitation, 300 μm coronal BLA slices were prepared in carbogenated ice-cold slicing solution using a vibrating-blade microtome (HM650V, Microm) equipped with a sapphire blade (Delaware Diamond Knives) and allowed to recover for 20 min at 33°C in carbogenated high-osmolarity artificial cerebrospinal fluid (high-Osm aCSF; [mM] 3.2 KCl, 11.8 glucose, 132 NaCl, 27.9 $NaHCO_3$, 1.34 NaH_2PO_4, 1.07 $MgCl_2$, 2.14 $CaCl_2$; 320 mOsm) followed by 40 min incubation at 33°C in carbogenated aCSF ([mM] 3 KCl, 11 glucose, 123 NaCl, 26 $NaHCO_3$, 1.25 NaH_2PO_4, 1 $MgCl_2$, 2 $CaCl_2$; 300 mOsm). Subsequently, slices were kept at RT in carbogenated aCSF until use. The submerged recording chamber was perfused with carbogenated aCSF at a rate of 2 ml · min^{-1} and maintained at 32°C.

Electrophysiological methods

Whole-cell patch clamp recordings were performed under visual control using an upright microscope (BX50WI, Olympus). Borosilicate glass pipettes (Harvard Apparatus 30-0068 Glass Capillaries GC150TF-7.5, 1.5 OD × 1.17 ID × 75 L mm) with resistances ranging from 3 to 7 megohms were pulled using a DMZ Universal electrode puller. Pipettes were filled using an intracellular solution allowing for EPSC and IPSC recording [(mM) 120 Cs-gluconate, 11 CsCl, 1 $MgCl_2$, 1 $CaCl_2$, 10 HEPES, 11 EGTA, 5 QX-314; 280 mOsm × kg^{-1}; pH adjusted to 7.3 with CsOH]. mCherry+ interneurons were visualized using epifluorescence and a 40× water immersion objective (LumPlanFl 40×/0.8, Olympus). Whole-cell patch clamp recordings were obtained from mCherry- putative principal LA/BLA neurons in regions showing robust somatic DREADD-mCherry expression. During voltage-clamp experiments neurons were held at 0 mV to measure IPSCs. Whole-cell voltage clamp recordings were performed using a MultiClamp 700B amplifier, filtered at 8 kHz and digitized at 20 kHz using a Digidata 1440A digitizer (Molecular Devices).

Electrophysiology data analysis

Data was analyzed using the Python package pyABF (https://swharden.com/pyabf). For line plots electrical stimulation responses were averaged in a 30 s sliding window. For quantification, responses 1.5 min to 0 min before CNO application were compared to responses 0.5 to 2.5 min after CNO application as indicated by the shaded area of the line plots to allow for CNO diffusion throughout the brain slice.

Calcium imaging data extraction

Basic image pre-processing: briefly, for each imaging session, recordings from all trials were concatenated into an image stack, corrected for bidirectional scanning (*55*), down sampled spatially by 2 by 2 pixels. Residual background was calculated with a disk-shaped morphological structuring element 'strel' in Matlab (MathWorks) and subtracted from the

signal. Then, images were corrected for rigid and for non-rigid motion using Normcorre (*55*). Registration across imaging sessions: after preprocessing as described, the mean intensity projections of each imaging session were registered using imregister in Matlab (MathWorks) and the resulting shifts applied to single frames. To check that GCaMP signal alone is sufficient for appropriate motion registration of subcellular compartments, we conducted a control experiment to compare image registration using either GCaMP or the florescence-independent fluorescent marker tdTomato as reference (fig. S2, H to M). For selecting subcellular compartments, non-overlapping somas, dendritic branches and dendritic spines were selected manually as regions of interest (ROIs) in ImageJ using the ROI manager tool (polygon selection for somas, line selection with 3 pixels ≈ 2 μm width for dendrites, oval selection with 4 to 7 pixels ≈ 3 to 5 μm for dendritic spines to account for small difference between imaging sessions) and carefully inspected for the absence of overlap with other subcellular compartments. The mean fluorescence intensity of all pixels from each ROI was considered for further analysis. Relative changes in Ca^{2+} fluorescence (F) were calculated using the formula: $\Delta F/F_0 = (F - F_0)/F_0$ (with F_0 = 33rd percentile of the fluorescence intensity over the entire trace) and used for all analysis of Ca^{2+} activity.

Data analysis

To identify responsive subcellular compartments, average Ca^{2+} traces were compared between the stimulus presentation period and a baseline period of the same duration. The stimulus response period was set to 1 s for shock presentations. For FC, US responsiveness was assessed by using the last second of CS presentation as baseline instead. Somas and dendrites with a CS offset response during the habituation day were not considered US responsive. If the mean Ca^{2+} activity during the stimulus period was larger or smaller than the mean Ca^{2+} activity during the baseline period by more than 2 SD (standard deviations), the ROI was classified as responsive (excited or inhibited, depending on response sign). Otherwise, the ROI was classified as non-responsive. To identify functional subclasses of neurons, we classified neurons in four mutually exclusive categories based on their CS responses on day 1 (habituation) and day 3 (test) as follows: (i) CS non-responsive: CS non-responsive on day 1 and day 3; (ii) CS neutral: same CS response sign (positive or negative response) on day 1 and on day 3 and absolute difference between the CS responses on day 1 and day 3 smaller than 15% ΔF/F0; (iii) CSup: CS response on day 3 larger than on day 1, excluding CS non-responsive and CS neutral neurons; (iv) CSdown: CS response on day 3 smaller than on day 1, excluding CS non-responsive and CS neutral neurons. For intra session CS response amplitude comparison, responses to the first and the last CS presentation on each session were compared. One-term Gaussian models was used to fit histogram distributions following the equation: $f(x) = a1*\exp\{-[(x - b1)/c1]^2\}$. All analysis was conducted in Matlab.

Probability and amplitude of stimulus response

To assess response probability, Ca^{2+} activity was compared between the stimulus presentation period and a baseline period of the same duration. The stimulus response period was set to 3 s for shock presentations. Response probability was defined as the proportion of trials where the mean Ca^{2+} activity during the stimulus period was larger than 2 SD (standard deviations) of the Ca^{2+} activity during the baseline period. Response amplitude was calculated from the average of all responsive trials. To compare the probability of response between days, the ratio of response probability was computed for each compartment. To compare the amplitude of response between days, the difference of response amplitude was computed for each compartment.

Calcium transient detection and normalization

Ca^{2+} transients were detected in somas, dendritic branches and dendritic spines using the findpeaks Matlab function with the following criteria: ΔF/F0 exceeding 1 SD (standard deviation) of the entire trace, prominence exceeding 0.5 SD (standard deviation) of the entire trace. This method reduces the influence of differences in signal-to-noise ratio when comparing Ca^{2+} activity between compartments (fig. S2). The same method was used to detect negative transients from the negative ΔF/F0 trace ($-\Delta F/F_0$) in order to estimate the rate of false-positive events (*56*). Within single neurons, Ca^{2+} transients in the dendrites and soma were defined as co-occurring when their peak amplitude was detected within a time window of 1 s in order to account for the variability of the Ca^{2+} indicator rise time in different compartments. Events were classified as compartment-specific (e.g., dendrite only) or co-occurring (e.g., soma + dendrite) relative to the dendrites or spines. For example, if co-occurring transients were detected in dendrite 1 and in its parent soma and undetected in dendrite 2, this transient was classified as soma + dendrite in dendrite 1 but not in dendrite 2).

Ca^{2+} transients detected during the tone presentation period were defined as tone-locked, Ca^{2+} transients detected outside of the tone presentation period as spontaneous. To compare the amplitude of Ca^{2+} transients in the soma and in dendritic branches, the Ca^{2+} trace of each compartment was normalized to its maximum $\Delta F/F_0$ over the course of all concatenated sessions. Correlation between somatic and dendritic Ca^{2+} activity was measured using the Pearson's correlation coefficient (r) of the normalized Ca^{2+} signal time course over the considered time period (*57*). Event amplitude correlation was measured with the Pearson's correlation coefficient (r) of the somatic and dendritic normalized amplitude of the automatically detected transients over the course of the entire considered session (5 transients minimum per session). Each session consists of concatenated stimuli presentations (30 s each, including 10 s pre-stimulus baseline). For spontaneous calcium events $\Delta F/F_0$ integral calculation and for comparison of spontaneous versus CS-locked $\Delta F/F_0$ integral, the area under the full width at half maximum of the automatically detected peak was considered. For calculation of CS-locked ΔF/F0 integral, the area under the curve during the tone presentation window (3 s) was considered.

Histology and fluorescent immunohistochemistry

Following completion of any behavioral paradigm and structural scans, mice were transcardially perfused. Mice were anesthetized with an intraperitoneal injection of an anesthetic mixture (250 mg/kg ketamine and 2.5 mg/kg medetomidine) to achieve deep anesthesia. The mice were then perfused transcardially with 0.1 M phosphate-buffered saline (PBS, pH 7.4) for 2 min followed by 4% paraformaldehyde (PFA) in PBS for 10-40 min. Then, the brains were removed from the skull and further incubated in in 4% PFA for 2-4 hours at 4°C, then cut into 100 μm coronal slices using a vibratome (VT1000S). Perfusion-fixed sections containing the amygdala were dried on slides and covered with Aqua-Poly/Mount (Polysciences). Sections were scanned with a laser scanning confocal microscope (LSM700) equipped with a 10× air objective (Plan-Apochromat 10×/0.45) or with a Zeiss Axio Scan Z1 Slide scanner equipped with a Zeiss objective (Fluar 5×/0.25) to confirm DREADD-mCherry expression and GRIN lens implantation site. GRIN lens placements were matched against a mouse brain atlas (fig. S1A) (*58*).

Statistical analyses and data presentation

Statistical analyses were carried out using Matlab (MathWorks). The sample sizes are similar to those used in the field. No statistical methods were used to determine sample size. Littermates were randomly allocated to experimental groups without pre-determined criteria and could be later identified by unique markers for group assignment. Experimenters were not blind to the group assignment of the animals. Animals were post hoc excluded from the analysis in the following three cases: (i) the

bottom tip of the GRIN lens was located outside the LA; (ii) brain motion was too great to be corrected post-hoc; (3) GCaMP6 labeling was not sparse enough to allow isolation of dendritic signals. The following number of mice were rejected, given for each per cohort: sensory stimuli tuning and stability: n = 1 mouse; FC: n = 4 mice; Unpaired conditioning: n = 3 mice, SST+ hM3D: n = 4 mice, SST+ mCherry: n = 2 mice, PV+ hM4D: n = 3 mice. The number of data points from separate cells and/or animal are indicated in figures, legends and supplementary tables. Averaging across multiple trials per cell/animal is indicated where applicable and n numbers always refer to data from individual cells/animals. Results described throughout the paper were reproduced. Multiple rounds of experimentation were required, i.e., from multiple mice, which were averaged for the presented datasets. Data was acquired from mice from multiple litters, and responses from individual cells were collected from at least four mice per group. All datasets were tested for normal distribution using a Shapiro-Wilk normality test. If the null hypothesis of normal distribution was not rejected, datasets were compared using a Student's t test or unpaired Student's t test, as appropriate. The corresponding one-sample tests were used to compare datasets to a fixed value. One-way analysis of variance (ANOVA) or Kruskal-Wallis tests were used when comparing more than two datasets, as appropriate. Post hoc multiple comparisons were performed using the Bonferroni correction. If the null hypothesis of normal distribution was rejected, datasets were compared using Wilcoxon signed-rank test or a Wilcoxon signed-rank test, as appropriate. The corresponding one-sample tests were used to compare datasets to a fixed value. A Fisher's exact test was conducted to compare proportions between two groups. Box-and-whisker plots show median values and 25th and 75th percentiles, the maximum whiskers length is 1.5 times the interquartile range. The corresponding mean ± SEM values for each test is reported in Table S1. A statistical significance threshold was set at 0.05, and significance levels are presented as $^{*}P < 0.05$, $^{**}P < 0.01$ or $^{***}P < 0.001$ in all figures. Averaging across multiple trials is indicated in the figure legends and respective methods sections where applicable. Contrast and brightness of representative example images were minimally adjusted using ImageJ or Zen lite Software (ZEISS technology). For figure display, Ca^{2+} traces were presented as $\Delta F/F_0$ or z-score ($z = (F - F_0) / \sigma$, with F_0 = mean fluorescence and σ = standard deviation over the entire Ca^{2+} traces).

REFERENCES AND NOTES

1. M. London, M. Häusser, Dendritic computation. *Annu. Rev. Neurosci.* **28**, 503–532 (2005). doi: 10.1146/annurev.neuro.28.061604.135703; pmid: 16033324
2. G. Major, M. E. Larkum, J. Schiller, Active properties of neocortical pyramidal neuron dendrites. *Annu. Rev. Neurosci.* **36**, 1–24 (2013). doi: 10.1146/annurev-neuro-062111-150343; pmid: 23841837
3. N.-L. Xu *et al*., Nonlinear dendritic integration of sensory and motor input during an active sensing task. *Nature* **492**, 247–251 (2012). doi: 10.1038/nature11601; pmid: 23143335
4. N. Takahashi, T. G. Oertner, P. Hegemann, M. E. Larkum, Active cortical dendrites modulate perception. *Science* **354**, 1587–1590 (2016). doi: 10.1126/science.aah6066; pmid: 28008068
5. J. Voigts, M. T. Harnett, Somatic and dendritic encoding of spatial variables in retrosplenial cortex differs during 2D navigation. *Neuron* **105**, 237–245.e4 (2020). doi: 10.1016/j.neuron.2019.10.016; pmid: 31759808
6. J. Cichon, W. B. Gan, Branch-specific dendritic Ca^{2+} spikes cause persistent synaptic plasticity. *Nature* **520**, 180–185 (2015). doi: 10.1038/nature14251; pmid: 25822789
7. M. E. J. Sheffield, D. A. Dombeck, Calcium transient prevalence across the dendritic arbour predicts place field properties. *Nature* **517**, 200–204 (2015). doi: 10.1038/nature13871; pmid: 25363782
8. A. Losonczy, J. K. Makara, J. C. Magee, Compartmentalized dendritic plasticity and input feature storage in neurons. *Nature* **452**, 436–441 (2008). doi: 10.1038/nature06725; pmid: 18368112
9. G. Schoenfeld *et al*., Dendritic integration of sensory and reward information facilitates learning. *bioRxiv* 474360 [Preprint] (2022). doi: 10.1101/2021.12.28.474360
10. L. Beaulieu-Laroche, E. H. S. Toloza, N. J. Brown, M. T. Harnett, Widespread and highly correlated somato-dendritic activity in cortical layer 5 neurons. *Neuron* **103**, 235–241.e4 (2019). doi: 10.1016/j.neuron.2019.05.014; pmid: 31178115
11. A. Kerlin *et al*., Functional clustering of dendritic activity during decision-making. *eLife* **8**, e46966 (2019). doi: 10.7554/eLife.46966; pmid: 31663507
12. V. Francioni, Z. Padamsey, N. L. Rochefort, High and asymmetric somato-dendritic coupling of V1 layer 5 neurons independent of visual stimulation and locomotion. *eLife* **8**, e49145 (2019). doi: 10.7554/eLife.49145; pmid: 31880536
13. M. S. Fanselow, A. M. Poulos, The neuroscience of mammalian associative learning. *Annu. Rev. Psychol.* **56**, 207–234 (2005). doi: 10.1146/annurev.psych.56.091103.070213; pmid: 15709934
14. H. T. Blair, G. E. Schafe, E. P. Bauer, S. M. Rodrigues, J. E. LeDoux, Synaptic plasticity in the lateral amygdala: A cellular hypothesis of fear conditioning. *Learn. Mem.* **8**, 229–242 (2001). doi: 10.1101/lm.30901; pmid: 11584069
15. C. W. Butler *et al*., Neurons Specifically activated by fear learning in lateral amygdala display increased synaptic strength. *eNeuro* **5**, 114 (2018). doi: 10.1523/ENEURO.0114-18.2018; pmid: 30027112
16. S. Maren, Neurobiology of Pavlovian fear conditioning. *Annu. Rev. Neurosci.* **24**, 897–931 (2001). doi: 10.1146/annurev.neuro.24.1.897; pmid: 11520922
17. M. T. Rogan, U. V. Stäubli, J. E. LeDoux, Fear conditioning induces associative long-term potentiation in the amygdala. *Nature* **390**, 604–607 (1997). doi: 10.1038/37601; pmid: 9403688
18. M. Bocchio, S. Nabavi, M. Capogna, Synaptic Plasticity, Engrams, and Network Oscillations in Amygdala Circuits for Storage and Retrieval of Emotional Memories. *Neuron* **94**, 731–743 (2017). doi: 10.1016/j.neuron.2017.03.022; pmid: 28521127
19. X. Zhang, B. Li, *Nat. Commun.* **9**, 1 (2018).doi: 10.1038/s41467-017-02088-w; pmid: 29317637
20. J. Gründemann *et al*., Amygdala ensembles encode behavioral states. *Science* **364**, eaav8736 (2019). doi: 10.1126/science.aav8736; pmid: 31000636
21. B. F. Grewe *et al*., Neural ensemble dynamics underlying a long-term associative memory. *Nature* **543**, 670–675 (2017). doi: 10.1038/nature21682; pmid: 28329757
22. J. P. Johansen, C. K. Cain, L. E. Ostroff, J. E. LeDoux, Molecular mechanisms of fear learning and memory. *Cell* **147**, 509–524 (2011). doi: 10.1016/j.cell.2011.10.009; pmid: 22036561
23. S. Krabbe *et al*., Adaptive disinhibitory gating by VIP interneurons permits associative learning. *Nat. Neurosci.* **22**, 1834–1843 (2019). doi: 10.1038/s41593-019-0508-y; pmid: 31636447
24. S. B. E. Wolff *et al*., Amygdala interneuron subtypes control fear learning through disinhibition. *Nature* **509**, 453–458 (2014). doi: 10.1038/nature13258; pmid: 24814341
25. C. Q. Chiu *et al*., Compartmentalization of GABAergic inhibition by dendritic spines. *Science* **340**, 759–762 (2013). doi: 10.1126/science.1234274; pmid: 23661763
26. Y. Kasugai *et al*., Structural and functional remodeling of amygdala GABAergic synapses in associative fear learning. *Neuron* **104**, 781–794.e4 (2019). doi: 10.1016/j.neuron.2019.08.013; pmid: 31543297
27. Y. Humeau *et al*., Dendritic spine heterogeneity determines afferent-specific Hebbian plasticity in the amygdala. *Neuron* **45**, 119–131 (2005). doi: 10.1016/j.neuron.2004.12.019; pmid: 15629707
28. S. Krabbe, J. Gründemann, A. Lüthi, Amygdala inhibitory circuits regulate associative fear conditioning. *Biol. Psychiatry* **83**, 800–809 (2018). doi: 10.1016/j.biopsych.2017.10.006; pmid: 29174478
29. J. J. Letzkus *et al*., A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* **480**, 331–335 (2011). doi: 10.1038/nature10674; pmid: 22158104
30. E. Abs *et al*., Learning-related plasticity in dendrite-targeting layer 1 interneurons. *Neuron* **100**, 684–699.e6 (2018). doi: 10.1016/j.neuron.2018.09.001; pmid: 30269988
31. G. Kastellakis, P. Poirazi, Synaptic clustering and memory formation. *Front. Mol. Neurosci.* **12**, 300 (2019). doi: 10.3389/fnmol.2019.00300; pmid: 31866824
32. M. Häusser, B. Mel, Dendrites: Bug or feature? *Curr. Opin. Neurobiol.* **13**, 372–383 (2003). doi: 10.1016/S0959-4388(03)00075-8; pmid: 12850223
33. G. Kastellakis, A. J. Silva, P. Poirazi, Linking memories across time via neuronal and dendritic overlaps in model neurons with active dendrites. *Cell Rep.* **17**, 1491–1504 (2016). doi: 10.1016/j.celrep.2016.10.015; pmid: 27806290
34. J. H. Choi *et al*., Interregional synaptic maps among engram cells underlie memory formation. *Science* **360**, 430–435 (2018). doi: 10.1126/science.aas9204; pmid: 29700265
35. D. W. Bloodgood, J. A. Sugam, A. Holmes, T. L. Kash, Fear extinction requires infralimbic cortex projections to the basolateral amygdala. *Transl. Psychiatry* **8**, 60 (2018). doi: 10.1038/s41398-018-0106-x; pmid: 29507292
36. H. Motanis, M. Maroun, E. Barkai, Learning-induced bidirectional plasticity of intrinsic neuronal excitability reflects the valence of the outcome. *Cereb. Cortex* **24**, 1075–1087 (2014). doi: 10.1093/cercor/bhs394; pmid: 23236201
37. M. Sehgal, V. L. Ehlers, J. R. Moyer Jr., Learning enhances intrinsic excitability in a subset of lateral amygdala neurons. *Learn. Mem.* **21**, 161–170 (2014). doi: 10.1101/lm.032730.113; pmid: 24554670
38. E. K. Lucas, A. M. Jegarl, H. Morishita, R. L. Clem, Multimodal and site-specific plasticity of amygdala parvalbumin interneurons after fear learning. *Neuron* **91**, 629–643 (2016). doi: 10.1016/j.neuron.2016.06.032; pmid: 27427462
39. R. D. Samson, É. C. Dumont, D. Paré, Feedback inhibition defines transverse processing modules in the lateral amygdala. *J. Neurosci.* **23**, 1966–1973 (2003). doi: 10.1523/JNEUROSCI.23-05-01966.2003; pmid: 12629202
40. P. Botta *et al*., Regulating anxiety with extrasynaptic inhibition. *Nat. Neurosci.* **18**, 1493–1500 (2015). doi: 10.1038/nn.4102; pmid: 26322928
41. A. G. Khan *et al*., Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. *Nat. Neurosci.* **21**, 851–859 (2018). doi: 10.1038/s41593-018-0143-z; pmid: 29786081
42. D. E. Wilson, B. Scholl, D. Fitzpatrick, Differential tuning of excitation and inhibition shapes direction selectivity in ferret visual cortex. *Nature* **560**, 97–101 (2018). doi: 10.1038/s41586-018-0354-1; pmid: 30046106
43. F. Zenke, W. Gerstner, S. Ganguli, The temporal paradox of Hebbian learning and homeostatic plasticity. *Curr. Opin. Neurobiol.* **43**, 166–176 (2017). doi: 10.1016/j.conb.2017.03.015; pmid: 28431369
44. K. A. Wilmes, H. Sprekeler, S. Schreiber, Inhibition as a binary switch for excitatory plasticity in pyramidal neurons. *PLOS Comput. Biol.* **12**, e1004768 (2016). doi: 10.1371/journal.pcbi.1004768; pmid: 27003565
45. A. P. Yiu *et al*., Neurons are recruited to a memory trace based on relative neuronal excitability immediately before training. *Neuron* **83**, 722–735 (2014). doi: 10.1016/j.neuron.2014.07.017; pmid: 25102562
46. P. Namburi *et al*., A circuit mechanism for differentiating positive and negative associations. *Nature* **520**, 675–678 (2015). doi: 10.1038/nature14366; pmid: 25925480
47. A. Beyeler *et al*., Organization of valence-encoding and projection-defined neurons in the basolateral amygdala. *Cell Rep.* **22**, 905–918 (2018). doi: 10.1016/j.celrep.2017.12.097; pmid: 29386133

48. P. Poirazi, A. Papoutsi, Illuminating dendritic function with computational models. *Nat. Rev. Neurosci.* **21**, 303–321 (2020). doi: 10.1038/s41583-020-0301-7; pmid: 32393820

ACKNOWLEDGMENTS

We thank all members of the Lüthi group for helpful discussions and comments on the work; G. Keller, D. Gerosa Erni, and T. Lu for virus production; P. Argast, C. Mueller, T. Eichlisberger, and all staff of the FMI Animal Facility for excellent technical assistance; J. Eglinger, S. Bourke, and R. Thierry at the FMI imaging facility and the FMI IT department for their dedicated technical support; and V. Jayaraman, R. Kerr, D. Kim, L. Looger, K. Svoboda, and the HHMI Janelia GENIE Project for making GCaMP6 available. **Funding:** This work was supported by the Novartis Research Foundation, the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (grant agreements 669582 to A.L. and 803870 to J.G.), the National Center of Competences in Research: "SYNAPSY—The Synaptic Bases of Mental Diseases" (financed by the Swiss National Science Foundation SNSF, 51NF40-158776), SNSF core (310030B_170268; all to A.L.), and SNSF Professorship (PP00P3_170672 to JG) grants; as well as a Young Investigator Grant from the Brain & Behavior Research Foundation (23593), a Career Development Award from the Synapsis Foundation–Alzheimer Research Switzerland ARS (2018-CDA02), and the Fondazione HED ALBIREO (to S.K.). A.S. was supported by a postdoctoral fellowship from the European Molecular Biology Organization (EMBO-ALTF-233-2020). **Author contributions:** S.D., S.K, J.G., and A.L. designed the project. S.D. and A.S. performed imaging experiments. S.D. performed imaging data analysis. M.M. performed in vitro patch-clamp experiments. S.D. and S.K. designed, generated, and tested the viral constructs. All authors contributed to the experimental design and interpretation of the data and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data and analyses necessary to understand and assess the conclusions of the manuscript are presented in the main text and in the supplementary materials. All processed data and scripts needed to evaluate the conclusions in the paper are available at https://data.fmi.ch/PublicationSupplementRepo with no access restriction.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abf7052
Figs. S1 to S11
Tables S1 and S2
References (*49*–*58*)
MDAR Reproducibility Checklist
Movies S1 and S2

13 November 2020; resubmitted 6 January 2022
Accepted 10 March 2022
10.1126/science.abf7052

Your Legacy to Science

AN ESTATE GIFT TO THE
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Since 1848, our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you become a member of our 1848 Society, joining Thomas Edison, Alexander Graham Bell and the many distinguished individuals whose vision led to the creation of AAAS and our world-renowned journal, *Science*, so many years ago.

Unlike many of its peers, *Science* is not for-profit. Your estate gift would provide long-term financial stability and durable annual income that will support operations and competitive innovation for years to come. **This support is vital.**

"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."

—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science
Federal Tax ID Number: 53-0196568
Address: 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

1848 SOCIETY AAAS

cut here

Yes, I would like more information about joining the AAAS 1848 Society.

PLEASE CONTACT ME AT:

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

RETURN THIS FORM TO:
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

RESEARCH ARTICLE

NEUROSCIENCE

Dynamic compartmental computations in tuft dendrites of layer 5 neurons during motor behavior

Yara Otor[1]†, Shay Achvat[1]†, Nathan Cermak[1], Hadas Benisty[2], Maisan Abboud[1], Omri Barak[1], Yitzhak Schiller[1], Alon Poleg-Polsky[3]*, Jackie Schiller[1]*

Tuft dendrites of layer 5 pyramidal neurons form specialized compartments important for motor learning and performance, yet their computational capabilities remain unclear. Structural-functional mapping of the tuft tree from the motor cortex during motor tasks revealed two morphologically distinct populations of layer 5 pyramidal tract neurons (PTNs) that exhibit specific tuft computational properties. Early bifurcating and large nexus PTNs showed marked tuft functional compartmentalization, representing different motor variable combinations within and between their two tuft hemi-trees. By contrast, late bifurcating and smaller nexus PTNs showed synchronous tuft activation. Dendritic structure and dynamic recruitment of the *N*-methyl-D-aspartate (NMDA)–spiking mechanism explained the differential compartmentalization patterns. Our findings support a morphologically dependent framework for motor computations, in which independent amplification units can be combinatorically recruited to represent different motor sequences within the same tree.

Cortical pyramidal neurons (PNs) typically possess an elaborate dendritic tree, which serves to receive and integrate the vast synaptic inputs arriving to the neuron. In vitro and modeling studies have established the role of passive, active, and morphological properties of the dendritic tree in information processing. Especially important are local dendritic *N*-methyl-D-aspartate (NMDA) and calcium spikes, which endow PNs with the capability of performing multicompartmental parallel nonlinear computations, potentially increasing the computational power and storage capacity of PNs and ultimately of the network (*1*–*5*).

In vivo evidence for local representation of information in different tuft dendrites of layer 5 PNs is scarce (*6*, *7*). Most reports show that the vast majority of calcium signals are highly correlated between different tuft branches, apical dendrites, and soma of same neurons (*8*–*13*), suggesting that the tuft and apical branches function mostly as a single compartment (*14*, *15*). Thus, current literature presents a puzzling gap between the powerful computational capabilities of these dendrites, as suggested by in vitro and modeling studies (*1*, *16*–*22*), and the much simpler computational scheme emerging from in vivo experiments.

In most in vivo studies, little consideration is given to the anatomical apical tuft structure of the layer 5 PNs. Layer 5 PNs are composed of two main anatomical subtypes that differ in their dendritic apical morphology: thick-tufted pyramidal tract (PT) and slender-tufted intra-telencephalic (IT) neurons (*23*–*30*). Thick-tufted PTNs are further subdivided into two distinct subtypes on the basis of their nexus morphology and molecular markers (*23*, *24*, *26*, *31*, *32*). Because the degree of compartmentalization and electrical coupling is strongly dependent on the dendritic morphology (*19*, *33*), we set out to examine motor representation in tuft dendrites of these two morphological types of thick-tufted layer 5 PTNs (*24*, *26*) in the primary motor cortex (M1) (*28*–*30*). We developed a comprehensive experimental and analysis platform to reveal the relationships between the detailed structural pattern and the behaviorally related calcium activity.

[1]Department of Physiology, Technion Medical School, Bat-Galim, Haifa 31096, Israel. [2]Yale University School of Medicine; Bethany, CT, USA. [3]Department of Physiology and Biophysics; University of Colorado School of Medicine, 12800 East 19th Avenue MS8307, Aurora, CO 8004, USA.
*Corresponding author. Email: jackie@technion.ac.il (J.S.); alon.poleg-polsky@cuanschutz.edu (A.P.-P.)
†These authors contributed equally to this work.

Results

We imaged the activity of tuft dendrites from single thick-tufted layer 5 PTNs in M1 forelimb cortex using two-photon calcium imaging during two head-fixed behavioral paradigms: a hand reach and grab for a food pellet (*34*) and running on treadmill tasks (*35*) (Fig. 1, A and B). We used the sparse labeling method of adeno-associated virus (AAV) viral vectors encoding two fluorescent proteins: GCaMP6 for recording the activity (*36*) and mRuby2 for reconstructing the morphology of the tree (Fig. 1, C to F) with a high degree of certainty. This resulted in the transfection of only a few layer 5 PTNs in our field of view (Fig. 1E; see materials and methods) and enabled us to trace tuft dendrites to their parent soma with high accuracy. Overall, we recorded 28 thick-tufted layer 5 PTNs in 22 mice. We first reconstructed the dendritic morphology from the two-photon volumetric Z-stack (Fig. 1C). We then used single-plane calcium imaging (30-Hz frame rate) to simultaneously record the activity of multiple dendritic tuft regions of interest (ROIs) of the same neuron (Fig. 1, E and F; on average, for each tree we sampled from 80.4 ± 11% of the terminal branches). These ROIs were aligned to the anatomical tree structure using custom software (see materials and methods).

Thick-tufted layer 5 PTNs are divided into two main well-established subtypes on the basis of their apical dendritic morphology with early and late bifurcating apical trunk (*24*, *26*, *27*). We used hierarchical clustering to subdivide our thick-tufted layer 5 PTNs on the basis of five morphological features (Fig. 1G and fig. S1, A to E). Consistent with previous studies, the clustering yielded two main subclasses: type 1, with early bifurcating apical tuft and long nexus, and type 2, with late bifurcating apical tuft and shorter nexus (Fig. 1G and fig. S1, F to H). Retrograde viral tracing of PTNs projecting to the medulla or cervical spinal cord revealed that M1 corticospinal PTNs yield a high proportion of type 1 dendritic morphology (fig. S1, I and J; 10 out of 11 of neurons in four mice were classified as type 1 neurons).

The two PTN subtypes demonstrated clear differences in their tuft calcium signals recorded during motor task performance. This was already apparent in the raw calcium imaging traces of individual trials (Fig. 1H, fig. S2, and movies S1 to S6). Type 1 thick-tufted layer 5 PTNs showed pronounced independent activity across their dendritic tuft branches (Fig. 1H and fig. S2), both in branches belonging to their right or left (R/L) tuft hemi-trees (Fig. 1H and fig. S2; division to left and right hemi-trees defined as all descendants of the first bifurcation branches) and within the hemi-trees (Fig. 1H). By contrast, type 2 PTNs did not show hemi-tree compartmentalization but did show more uniform dendritic tuft activation across their entire tuft tree (Fig. 1H and fig. S2).

To quantitatively investigate the correlations between the detailed tuft tree structure and dendritic activity in type 1 and 2 PTNs, we constructed and compared ROI distance matrices on the basis of dendritic structure and calcium activity. We obtained a structural distance matrix of the tuft tree by measuring the shortest path distance between all pairwise combinations of ROI locations, organized according to the tree structure (Fig. 1I). Calcium activity was extracted by identifying the calcium events in each ROI for all trials with the modified event detector MLSPIKE (*37*) (materials and methods and fig. S3A).

Calcium event amplitudes exhibited a long-tailed asymmetrical distribution (Fig. 1J and

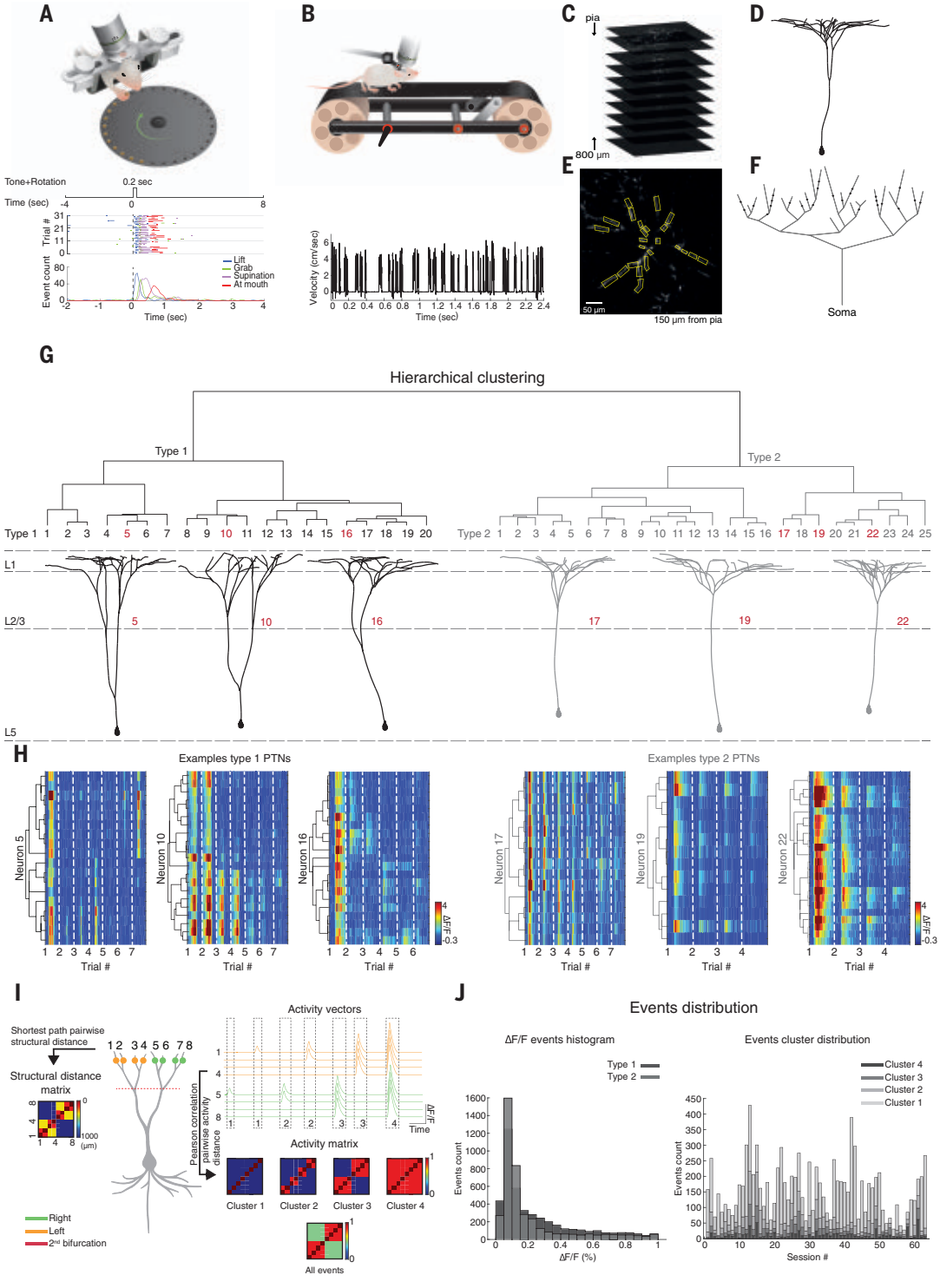


Fig. 1. M1 contains two populations of thick-tufted layer-5 PTNs with distinct morphologies and functional properties. (**A**) (Top) Behavioral setup for hand reach task. (Bottom) An ethogram and histogram annotating the behavior of an expert mouse. (**B**) (Top) Behavioral setup of head-fixed mice running on a treadmill. (Bottom) Example of velocity (cm/s) over time (s). (**C** and **D**) Z-stack of a single PTN acquired in vivo and its three-dimensional anatomical reconstruction. (**E**) Single-imaging plane of a GCaMP6s-positive layer 5 PTN's tuft dendrites. ROIs are marked in yellow. (**F**) Schematics of the tuft morphology; dots represent ROIs. Same neuron in (C) to (F). (**G**) (Top) Hierarchical clustering of thick-tufted layer 5 PTNs based on five morphological parameters. (Bottom) Three example apical morphologies from each PTN type. Basal and oblique dendrites are not shown. Red numbers indicate the corresponding neurons in the clustering. (**H**) Calcium events ($\Delta F/F$ heat map) in single trials from the same three type 1 (left) and type 2 (right) PTNs shown in (G), arranged by the tree structure as indicated by the dendrogram. (**I**) Illustration of structure-activity correlations. Schematic representation of a layer 5 PTN with structural distance matrix presented as a heat map of the shortest path distance between each two ROIs (left) and illustrated activity traces for ROIs (right upper panel, dashed rectangles mark events belonging to different event clusters) and activity matrix showing Pearson correlation between ROI pairs presented for all events and for each event cluster (right lower panel). (**J**) (Left) Normalized calcium events histogram from 21 neurons for type 1 (black) and type 2 (gray) PTNs. (Right) Event distribution color coded by event cluster type for individual sessions in multiple neurons.

Fig. 2A). This raised the possibility of multiple subpopulations of calcium events. Indeed, visual inspection suggested four qualitatively distinct types of activation patterns in type 1 PTNs (single branch, subtree, hemi-tree, whole tree; Fig. 1, H and I; Fig. 2B; and fig. S2). To perform quantitative analysis of these different event types, we subdivided the events into four clusters according to either their average amplitudes over all ROIs (see materials and methods) or the number of activated ROIs during each event (spatial activation extent) using the k-means algorithm (average percentage of events in cluster 1 = 66%, cluster 2 = 17%, cluster 3 = 10%, cluster 4 = 8% of the events; Fig. 1H, Fig. 2B, and figs. S2 and S3). With the calcium activity vectors and the structural distance matrix at hand, we constructed a pairwise Pearson correlation activity matrix arranged according to the structural distance matrix for each of the

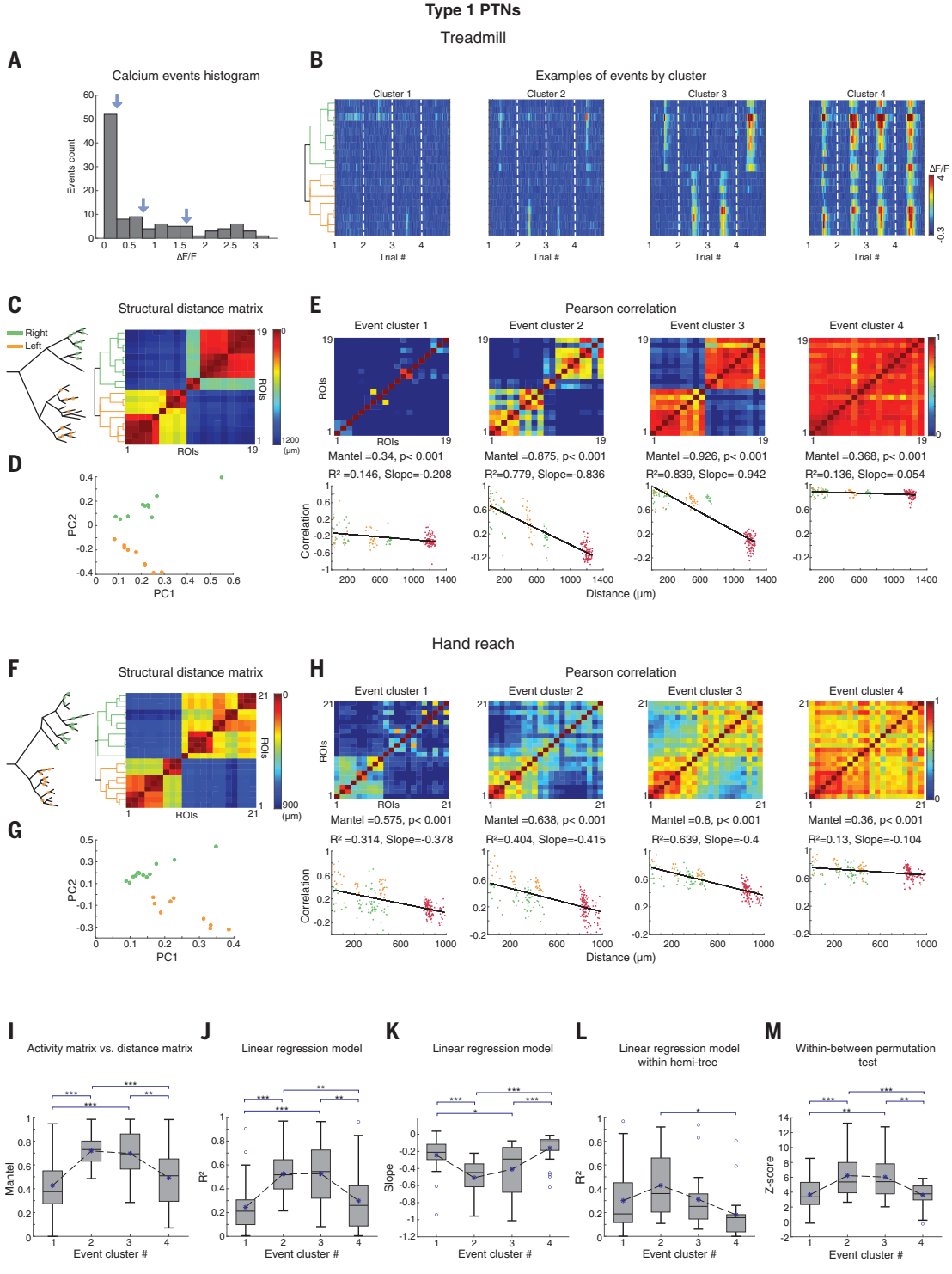


Fig. 2. Compartmentalized activity in tuft dendrites of type 1 thick-tufted layer 5 PTNs during running on treadmill and hand reach. (**A**) Calcium events histogram from one type 1 PTN during treadmill session. Arrows indicate values separating the four different event clusters. (**B**) Examples of calcium events ($\Delta F/F$ heatmap) from the four event clusters in type 1 PTN. ROIs are arranged by the tree structure as indicated by the dendrogram, shown on the left (R, green; L, orange). (**C**) Two-dimensional tree diagram (left) and the corresponding structural distance matrix and dendrogram (right) of a type 1 PTN. Dots represent recorded ROIs. (**D**) Two-dimensional PCA embedding of all ROIs activity; each dot represents a single ROI. (**E**) (Top) Matrices showing pairwise Pearson correlation coefficients computed from the calcium signals arranged by the tree structure shown in (C). (Bottom) Pearson correlation values as a function of shortest path distance fitted with a linear regression model. [(A) to (E)] Same neuron and session. ROIs compared within left hemi-tree (orange); within right hemi-tree (green) and between R/L hemi-trees (red). Black line represents linear regression model fit. (**F** to **H**) As in (C) to (E) for a different type 1 PTN during a hand reach session. (**I** to **M**) Box plots of the following parameters: Mantel statistics comparing Pearson correlation and structural distance matrices (I), R^2 of linear regression model that predicted Pearson correlations by distance (J), slope of linear regression model that predicted Pearson correlations by distance (K), R^2 of linear regression model that predicted Pearson correlations by distance calculated for the hemi-trees separately (L), and Z-score of the experimental Pearson correlations of within compared to between hemi-trees (mean $\text{Pearson}_{\text{within}}$ − mean $\text{Pearson}_{\text{between}}$) calculated in relation to the shuffled distribution (M). *p < 0.05, **p < 0.01, ***p < 0.001; blue asterisk, mean value. One-way analysis of variance (ANOVA) with Tukey post-hoc test (12 neurons, 8 animals, 27 sessions).

four calcium event clusters and for all events (Fig. 1, I and J).

We quantified the anatomical-functional correlations using three measures: First, we performed principal component analysis (PCA) to embed the activity of each ROI across time and trials into a two-dimensional space of the two leading components that explain 84.42 ± 9.67% of the variance. Second, we used a Mantel test (*38*–*40*), which quantifies the degree of correlation between the functional activity and structural distance matrices. Third, we

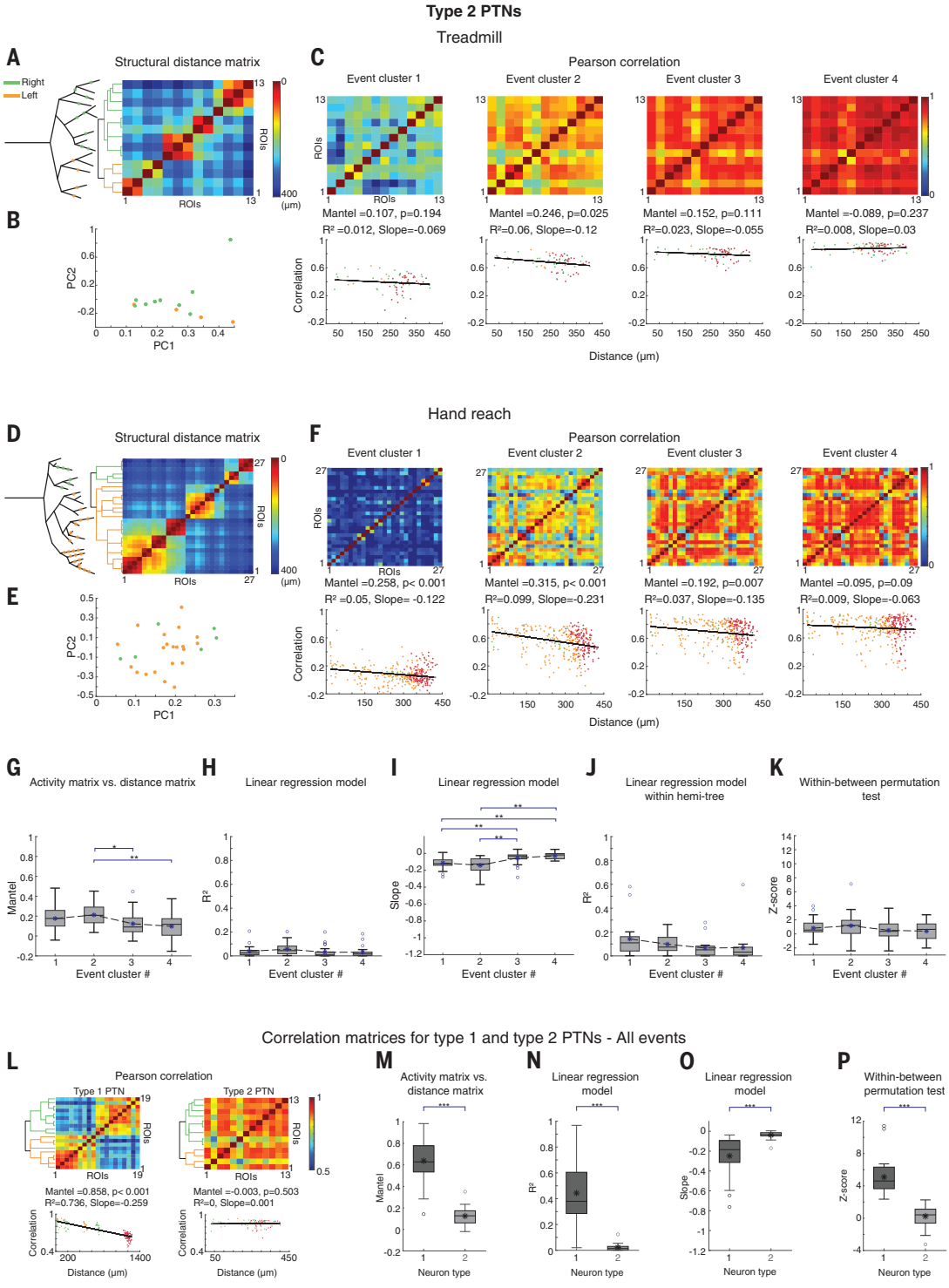


Fig. 3. Homogeneous activity in tuft dendrites of type 2 thick-tufted layer 5 PTNs during running on treadmill and hand reach. (**A** to **K**) As in (C) to (M) of Fig. 2 for type 2 PTNs. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; blue asterisk, mean value. One-way ANOVA with Tukey post-hoc test (9 neurons, 9 animals, 25 sessions). (**L**) Pairwise Pearson correlation coefficient matrices computed for all calcium events for type 1 (left) and type 2 (right) PTNs. (Bottom) Pearson correlation values as a function of shortest path distance fitted with a linear regression model for all events calculated for type 1 (left) and type 2 (right) PTNs. Type 1 PTN as in Fig. 2, C to E, and type 2 as in (A) to (C). (**M** to **P**) Box plots are the same as in (G) to (I) and (K) when all calcium events are considered for type 1 and type 2 PTNs.

measured the relationship between the pairwise Pearson activity correlations to the distance between ROIs with a linear regression (Fig. 2, C to H).

We first concentrated on type 1 PTNs. We observed strong compartmentalization of the tuft tree with a significant correlation between the tree structure and calcium activity during both motor tasks: hand reach and running on treadmill. The first cluster, with the smallest calcium event amplitudes, typically involved correlated activity of individual or sister branches (Fig. 2, E and H). For this cluster, the Mantel statistics were relatively low (Fig. 2I), as was the linear correlation between the Pearson values and dendritic distance (Fig. 2, J to L, and fig. S4A; linear regression permutation test). For events with intermediate amplitudes corresponding to clusters 2 and 3, we observed high correlation between the activity and the tree structure, with strong compartmentalization of the tuft tree, ranging from correlated activity mapped to secondary and tertiary branches (event cluster 2; Fig. 2, C to H) or entire hemi-trees (event cluster 3; Fig. 2, C to H). The strong correlation between the anatomical and activity matrices in these

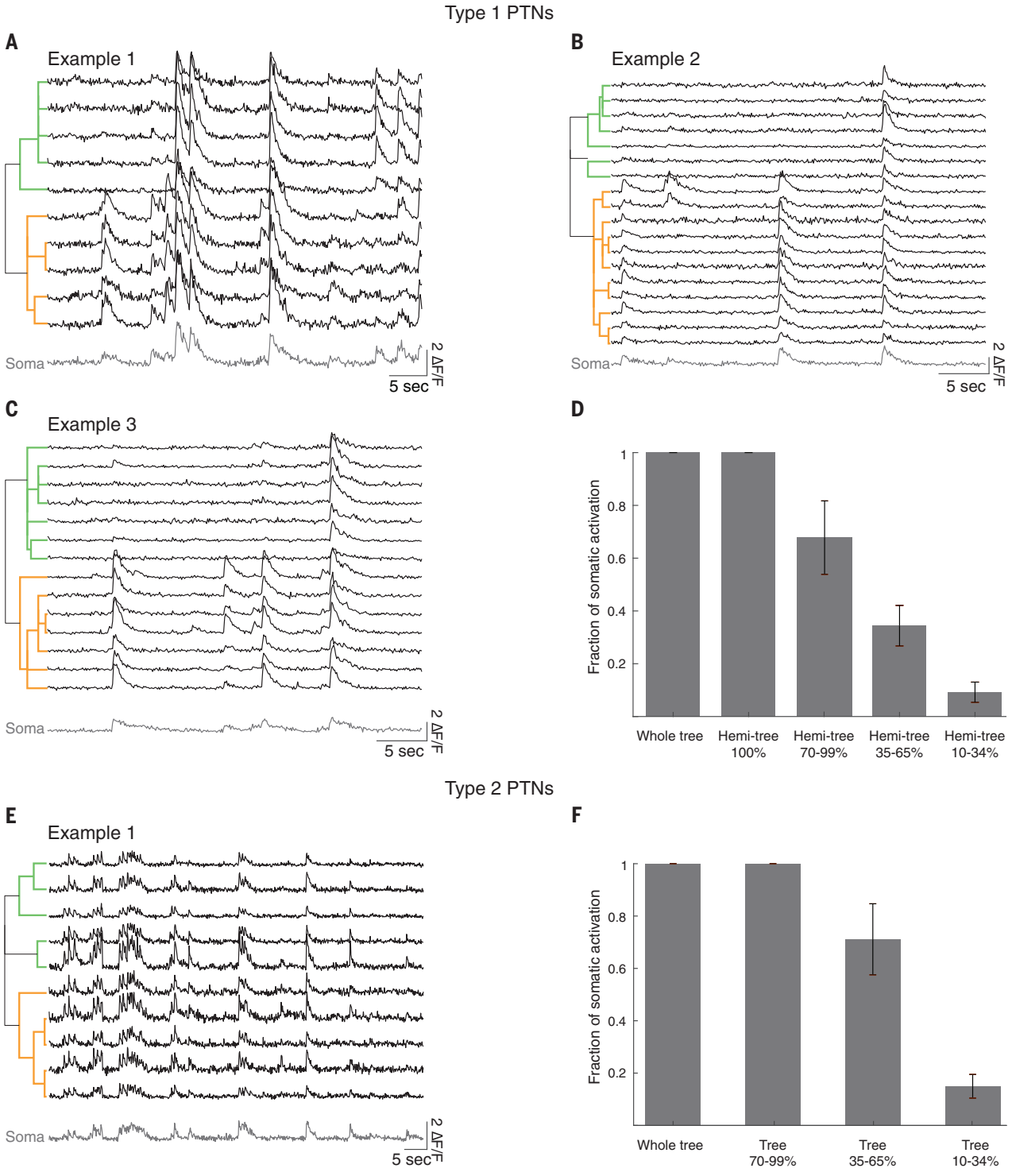


Fig. 4. Tuft and soma activity correlations. Quasi-simultaneous tuft and soma recordings were performed with an electrically tunable lens (ETL) (~10 Hz) in type 1 and 2 PTNs. (**A** to **C**) Examples of representative Δ*F*/*F* traces in the tuft and soma from a type 1 PTN (ROIs are arranged by the tree structure as indicated by the dendrogram). (**D**) Mean (±SD) fraction of tuft events that were simultaneously active with the soma for type 1 PTNs (seven neurons, six animals). (**E**) As in (A) for a type 2 PTN. (**F**) As in (D) for type 2 PTNs (three neurons, three animals).

clusters resulted in higher Mantel statistics (Fig. 2I) and a stronger dependence between pairwise Pearson correlations and tuft dendritic distance (Fig. 2, E, H, and J to L, and fig. S4A; linear regression permutation test). Finally, for the largest-amplitude calcium events (cluster 4), the activity was spread throughout the entire tuft tree, and responses in all ROIs were highly correlated (Fig. 2, E and H). This was manifested as lower Mantel statistics (Fig. 2I) and lower dependence of the pairwise Pearson correlation values on dendritic distance (Fig. 2, E, H, and J to L, and fig. S4A; linear regression permutation test). Even in cluster 4 events, it is evident that some

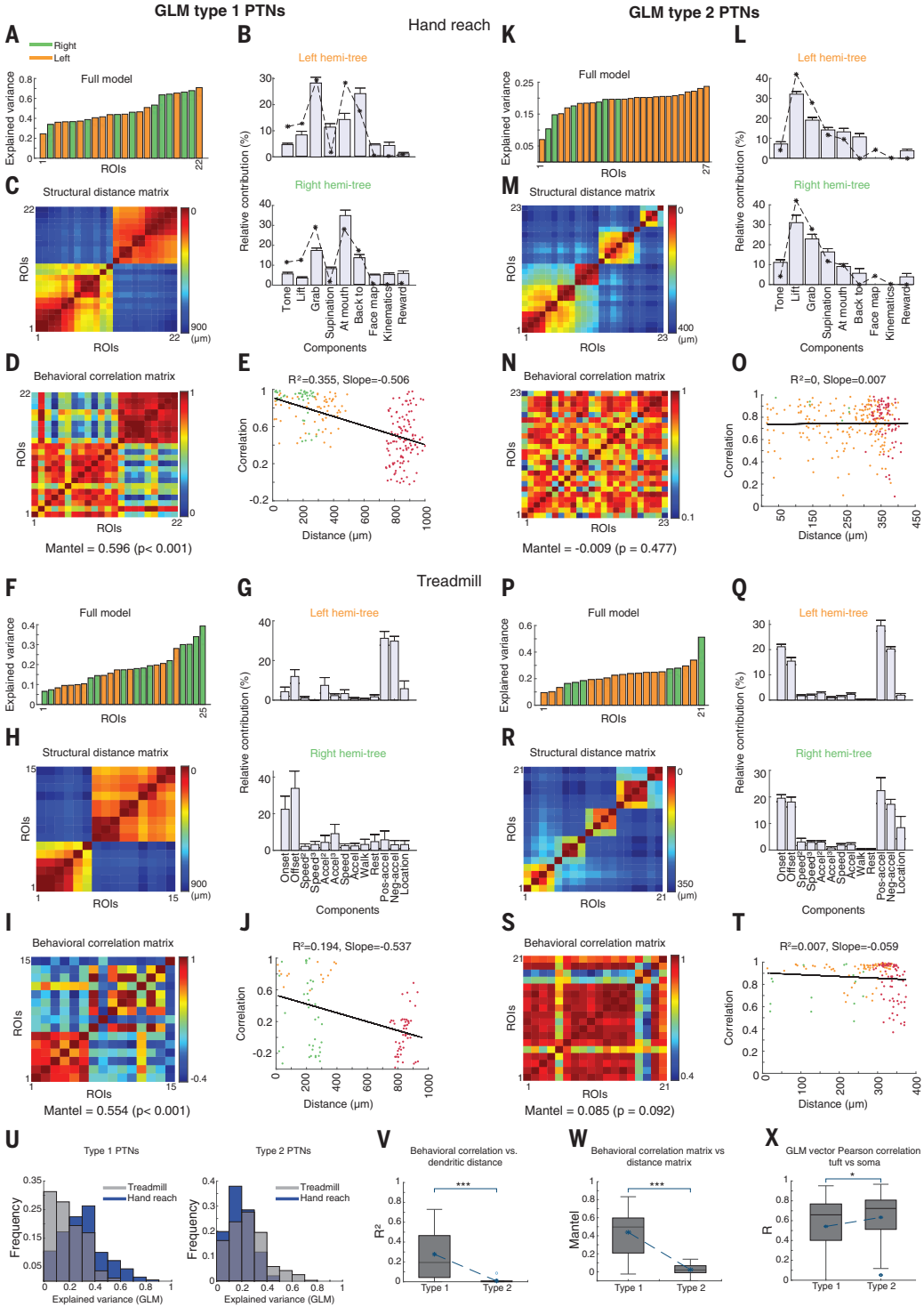


Fig. 5. Modeling the contribution of the behavioral predictors to calcium activity of type 1 and type 2 PTNs using a generalized linear model (GLM). (**A**) Goodness of fit (R^2) of GLM full model for each ROI from a type 1 PTN during a hand reach session. (**B**) Mean relative contribution of modeled ROIs (R^2 >15%) for each of the behavioral predictors for right and left hemi-trees. Black asterisks on both graphs indicate values computed from somatic recordings of same neuron. (**C**) Structural distance matrix. ROIs with R^2 < 0.15 were excluded. (**D**) Pairwise Pearson correlation between the GLM relative contribution vectors for all included ROI pairs, arranged by the tree structure. (**E**) Pearson correlation values shown in (D) as a function of shortest path distance fitted with a linear regression model. (**F** to **J**) As in (A) to (E) during running on a treadmill. Same neuron in (A) to (J). (**K** to **T**) As in (A) to (J) for an example type 2 PTN. (**U**) (Left) Frequency of ROI's R^2 values of GLM full model for type 1 PTNs (10 animals, 14 neurons, 31 sessions). (Right) As in left panel for type 2 PTNs (9 animals, 10 neurons, 31 sessions). (**V** to **X**) Box plots of the following parameters: R^2 linear regression model that predicted the Pearson correlations of GLM relative contribution vector by dendritic distance (V), Mantel statistics comparing the structural distance matrix and the behavioral-correlation matrix (W), Pearson correlation between the soma's behavioral relative contributions and those of the tuft, for type 1 and type 2 PTNs (seven and four neurons, respectively) (X). ***p < 0.001; blue asterisks, mean value. Wilcoxon rank test.

portions of the tree were more strongly correlated than others. Similar results were obtained when event clustering was based on the spatial extent of calcium signals instead of their averaged amplitude (fig. S5).

To further examine the R/L hemi-tree segregation, we used two additional analysis methods: First, we shuffled the "within" and "between" tagging of ROI Pearson correlation pairs while maintaining their pairwise correlation coefficient values. We found the experimental values of the difference between the mean Pearson correlations of within compared to between hemi-trees to be significantly higher compared to the shuffled distribution for all cluster events (Fig. 2M and fig. S6, A and C; Z-score >3.6). Second, we calculated the proportion of the variance explained (R^2) of the calcium activity of each ROI in one hemi-tree by the activity of all ROIs in the contralateral hemi-tree using a linear regression model. The results further indicated compartmentalization of activity in the hemi-trees in clusters 1 to 3 (fig. S7A).

To facilitate the comparison between type 1 and type 2 PTNs, we also divided the events of type 2 PTNs into four clusters. The distributions

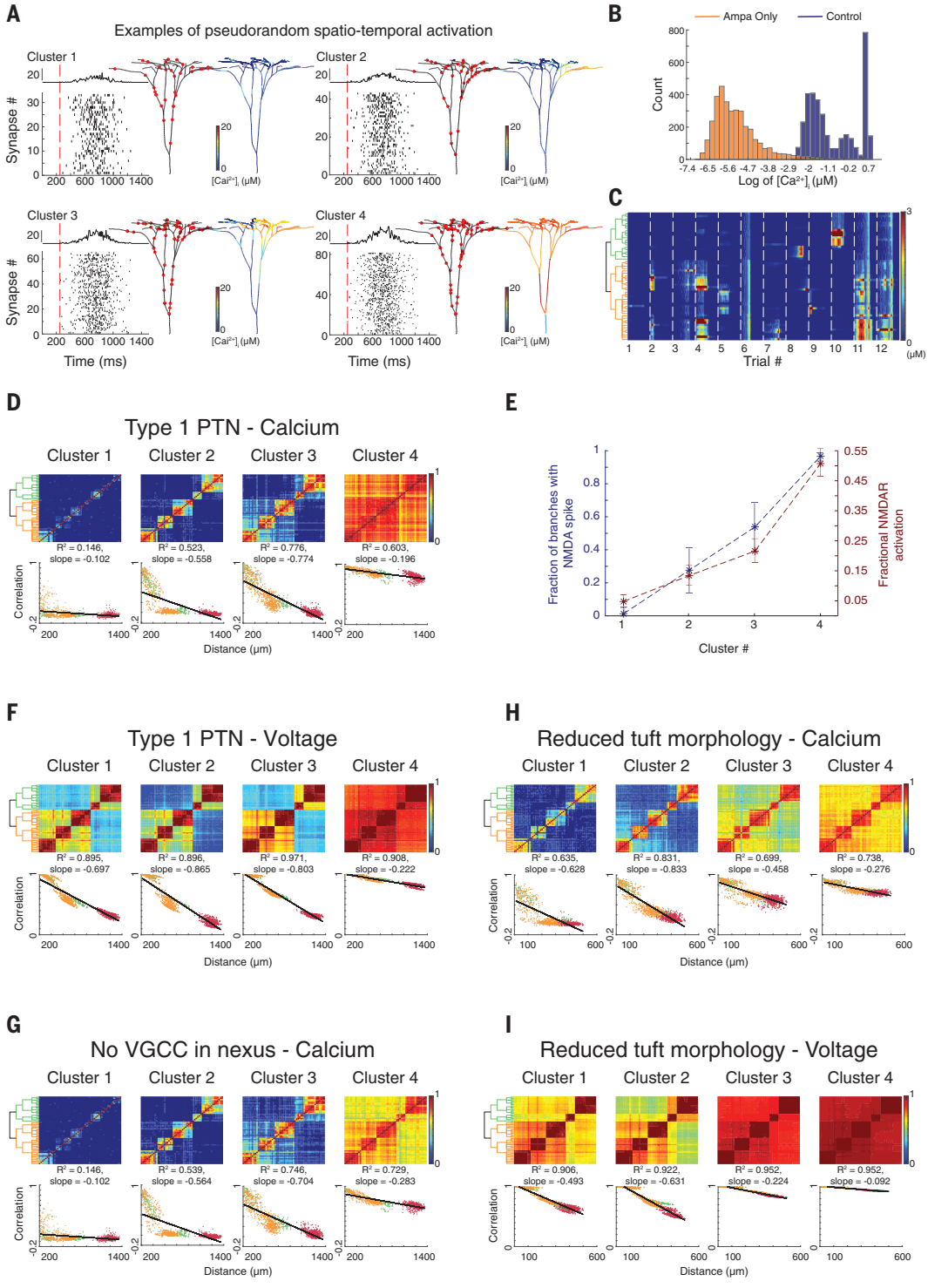


Fig. 6. Simulation of type 1 PTN explains in vivo activity by apical morphology and NMDA spikes. (**A**) Examples of the temporal distribution of simulated pseudo-random synaptic activation patterns for the four event clusters in one trial. Top trace, the total number of activated synapses over time. (Right) The corresponding spatial input distribution and post-synaptic calcium signal (shown on a logarithmic scale). The average number of recruited synapses for each event cluster was 34 ± 13, 48 ± 13, 60 ± 15, and 84 ± 29. (**B**) The distribution of the simulated evoked calcium events in the tuft dendrites of the reconstructed type 1 PTN. (**C**) Representative calcium activity in different simulation trials, arranged by the tree structure as indicated by the dendrogram, left. (**D**) (Top) Pairwise Pearson correlation coefficients computed from the tuft calcium signals arranged by the tree structure. (Bottom) Pearson correlation values as a function of shortest path distance fitted with a linear regression model (black). (**E**) The normalized number of tuft dendrites with NMDA spikes (blue) and the fractional NMDAR conductance (brown) for all event clusters. Error bars: SD. (**F**) As in (D), for tuft voltage correlations. (**G**) As in (D), in the absence of VGCC in the nexus. (**H**) As in (D), when the morphology of the tuft was reduced to match the extent of the tuft dendrites in type 2 PTNs shown in fig. S13A. (**I**) As in (H), for tuft voltage correlations. Color coding: orange and green represent data from the left and right hemi-trees, respectively. Red, comparison between R/L hemi-trees. Simulations for this figure are for the neuron in Fig. 2, C to E.

of peak calcium events and the peak calcium amplitudes of the clusters were comparable in type 1 and 2 PTNs (fig. S3, C and D). Type 2 PTNs did not show significant correlations between their tree structure and functional calcium activity for any of the calcium event clusters (Fig. 3 and figs. S4B and S7, B to E), and compartmentalization between their R/L hemi-tree was low (figs. S6, B and C, and S7A). Thus, for type 2 PTNs except cluster 1, calcium events globally involved the entire tuft tree, consistent with previous reports in visual, anterior lateral motor (ALM), and somatosensory cortices (*8*–*11*). The differences in the structure-function correlations between type 1 and type 2 PTNs were not related to the cluster subdivision. They were also evident when the analysis was performed on the whole event population (Fig. 3, L to P). However, in type 1 PTNs, the subdivision into the four clusters highlighted more details because using the entire population averaged events with different spatial activation patterns.

Plotting Mantel statistics as a function of nexus size for individual neurons, we found a clear distinction between the two PTN subclasses despite the variability in size within both groups Support Vector Machine (SVM) accuracy 1,

chance 0.52 for both clusters). This finding further strengthens the anatomical subdivision based on independent physiological parameters (fig. S7F).

We next investigated the relationship between tuft tree and somatic activation using quasi-simultaneous imaging of tuft and soma (~10-Hz acquisition rate). For type 1 PTNs, we observed that events encompassing the full tuft or restricted to an entire hemi-tree were invariantly associated with somatic activation in all (100%) events examined. For events encompassing only part of a hemi-tree, somatic activation was proportional to the percentage of active ROIs (Fig. 4, A to D). In type 2 PTNs, somatic activation was proportional to the extent of ROI recruitment in the entire tree (Fig. 4, E and F). The tuft activity was critically dependent on NMDA receptor (NMDR) channels because local injection of the NMDAR blocker MK801 close to the imaged tuft blocked both tuft and somatic activity (fig. S8).

Next, we investigated the possible functional importance of tuft compartmentalization. For both type 1 and type 2 PTNs, the activity in tuft dendrites strongly correlated with motor behavior (fig. S9). To evaluate the preferential selectivity of responses for specific behavioral variables in the different tuft dendrites in individual neurons, we modeled the calcium transients using a generalized linear model (GLM) (Fig. 5 and fig. S10A; see materials and methods) (*34*, *41*, *42*). For both type 1 and type 2 layer 5 PTNs, the activity in tuft dendrites was strongly related to individual motor variables. The GLM effectively modeled the calcium activity of both hand reach and running on treadmill behavioral events (Fig. 5). On average, for type 1 and type 2 PTNs, the full GLM successfully modeled the ROIs activity, achieving explained variance >0.15 in 77.5 and 63.8% of ROIs for the hand reach task and 56.5 and 73% of ROIs for the treadmill task, respectively (Fig. 5U).

For type 1 PTNs, representation was not uniform throughout the tuft for either hand reach or treadmill behaviors. We observed a differential representation of motor variables in the different tuft tree segments (Fig. 5, A to J and V to W, and fig. S10B). The largest nonuniformity was typically observed between the R/L tuft hemi-trees with different combinations of motor variables preferentially encoded by each of the two hemi-trees activity (Fig. 5, A to J, and fig. S10B). However, we also could observe dendrites within each hemi-tree, which were tuned to different combinations of motor variables (fig. S10B). To quantify the spatial compartmentalization of motor variables representation within and between the hemi-trees of single type 1 PTNs, we performed pairwise Pearson correlations between the GLM selectivity vectors of the different ROIs (Fig. 5, D, E, I, and J, and fig. S10B). Overall, type 1 PTNs demonstrated a correlation between the pairwise Pearson correlation coefficient of the GLM selectivity vectors and the distance between ROIs (Fig. 5, A to J and V). Consistently, we observed a significant correlation between the tuft distance matrix and GLM vector matrix for both behavioral data (Fig. 5, V and W). The significance of R/L segregation was further examined by comparing our experimental Pearson correlations between R/L hemi-trees to randomly distributed ROIs (1000 permutations). Experimental Pearson values between R/L hemi-trees were significantly smaller compared with the permutated values (Z-scores of −2.94 ± 2.6; p values were <0.05 in 80% of cases). Moreover, when we plotted the Pearson correlation matrices of peri-behavioral event time segments for different behavioral events, we found marked differences in the dendritic compartmentalization between different behavioral events for both tasks (fig. S11), further indicating differential spatial dendritic representation of various motor events.

By contrast, GLM analysis for type 2 PTNs revealed a more uniform encoding of motor variables along the tuft and only minimal correlation to the tuft tree structure (Fig. 5, K to T). Although the GLM vectors for the different ROIs of the same single neuron were not completely identical, we did not observe systematic differences between or within the hemi-trees (Fig. 5, V and W; comparison of experimental and permutated random ROI locations yielded Z-scores of −0.47 ± 1.72; p values were <0.05 in 27% of cases; $p < 0.01$ comparing between Z-scores of type 1 and type 2 PTNs, Wilcoxon rank test).

To understand the impact of the differential tuft activity on the output of type 1 and type 2 PTNs, we performed tuft and soma imaging from the same neurons. We hypothesized that the combined computational products of the hemi-trees will affect the representation of motor variables at the soma. Indeed, GLM modeling of the somatic activity revealed a composite tuning that reflected the summed representation of both hemi-trees for type 1 and 2 PTNs (Fig. 5, B and L, and fig. S12). We found high Pearson correlation coefficients between the GLM vectors of the R/L tuft ROIs and the corresponding somas for type 1 and even more so for type 2 PTNs (Fig. 5X).

To further investigate the mechanisms underlying the observed activity in vivo, we performed modeling experiments on reconstructed type 1 and type 2 PTNs using the neuron platform (Fig. 6). We activated the tuft tree with pseudorandom patterned inputs (Fig. 6, A to C) and in vivo–like activation frequencies (*43*, *44*). Compatible with our in vivo results, we could readily observe R/L hemi-tree separation and compartmental calcium responses within hemi-tree tuft branches in a type 1 (Fig. 6C; compare Fig. 6D and Fig. 2, C to E, and movies S7 to S10; same neuron as in Fig. 2, A to E) but not in type 2 PTNs (fig. S13; same neuron as in Fig. 3, A to C). However, the in vivo results showed a higher intercorrelation within terminal subtrees for clusters 2 and 3 events, probably reflecting a nonrandom input distribution on the tuft tree in vivo (fig. S14, A and B). The simulated voltage correlation matrices also captured the main features of the tuft tree segmentation (Fig. 6F). Yet, the voltage signals showed higher correlations than the calcium signals, reflecting the localized calcium influx to the synaptic sites (via NMDAR) and the high-pass filtering effect of voltage-gated calcium channels (VGCCs) [see also Lavzin *et al.* (*45*)].

Next, we used our simulations to characterize the nature of the calcium event clusters. In type 1 PTNs, the four clusters differed primarily in the degree of recruitment of the NMDAR spiking mechanism (Fig. 6E). In type 2 PTNs, massive and widespread recruitment of NMDA spikes already occurred in cluster 2 events (fig. S13, D and H). Consistent with the critical role of NMDAR spiking mechanisms in replicating the in vivo findings, simulations with AMPAR-only synapses (Fig. 6B) failed to generate significant tuft activation (*46*). These results agree with our experimental demonstration of a marked reduction in the calcium activity of tuft dendrites after NMDAR blockade (fig. S8). We then investigated the contribution of VGCCs to tuft compartmentalization in type 1 PTNs. Elimination of VGCCs from the nexus had little impact on tuft segmentation, except for cluster 4 events, where removal of VGCCs increased R/L hemi-tree segmentation (Fig. 6G). Taken together, our simulations show that tuft segmentation is primarily dependent on differential recruitment of NMDAR spiking mechanisms in different tuft branches, and VGCCs play a minimal role in this process.

Because we could replicate the major characteristics of our experimental results with pseudo-random input patterns in both type 1 and 2 PTNs, we hypothesized that the distinct degrees of tuft segmentation must be linked to the difference in dendritic morphology, especially nexus size. To investigate this hypothesis, we scaled down the size of the apical arbor of the type 1 PTN but kept the soma-tuft distance unchanged. Under these conditions, we observed a sharp reduction in the segmentation within and between the hemi-trees (Fig. 6, H and I). This conclusion is further supported by our voltage simulations, which show that dendritic independence was primarily driven by the sizable electrotonic distance along the nexus (fig. S14C) and by the decreased NMDAR-dependent nonlinear interactions as a function

of the distance between dendritic locations, with only minimal interactions between the two hemi-trees [fig. S14D; see also Kerlin *et al.* (*11*)].

Discussion

Overall, our results reveal a subclass of thick-tufted layer 5 PTNs in M1 (type 1) that performs parallel independent representations of motor information within its tuft dendrites. The degree of tuft compartmentalization was primarily dependent on the nexus and tuft tree morphology and required the NMDA spiking mechanism. In these neurons, motor information is integrated and amplified via NMDA spikes within adjustable segments of the tuft, ranging from small tertiary and quarterly sister branches, independent R/L hemi-trees, and up to a single global computational compartment in the minority of events. Our modeling suggests that calcium spikes play only a minimal role in tuft compartmentalization and local amplification. Instead, the initiation of calcium spikes in the distal apical trunk and proximal nexus branches probably serves to further amplify the tuft computational products (*46*, *47*).

Type 2 PTNs also amplify behaviorally relevant motor information within their tuft but in a more global manner, mostly functioning as a single computational compartment (*8*–*11*).

The type 1 subclass constitutes a sizable fraction of thick-tufted layer 5 PTNs (~40%), in line with previous reports (*24*, *26*). Although we based our classification solely on the apical dendritic morphology, our physiological findings supported this classification. Studies that examined both dendritic morphology and projection targets suggest that type 1 PTNs may preferentially project to the medulla (*23*, *27*, *32*), which is consistent with our medulla and spinal cord retrograde-labeled PTNs. Further studies are required to examine the projection pattern and molecular markers of type 1 and 2 PTNs (*23*, *32*).

Our results reconcile the differences in the findings of prior in vitro and modeling studies, which predicted the capacity of dendrites to compartmentalize information (*4*, *17*, *22*, *48*, *49*) and recent recordings from behaving mice, which show that tuft dendrites function primarily as a single global amplification unit (*8*–*11*). Several past studies also observed infrequent local spikes that were limited to small, non-overlapping dendritic segments in tuft dendrites (*9*–*11*). Yet, this highly localized spiking activity, reminiscent of our cluster 1 events, cannot serve for efficiently communicating tuft computations to the soma. It may be used for local plasticity instead (*6*, *21*). A study (*6*) that recorded tuft dendrites of M1 layer 5 PTNs reported spatially isolated dendritic spikes in almost all pairs of sibling distal tuft branches (95%). The results of this study differ from ours, as we observed selective activation of sibling terminal tuft branches during both motor tasks infrequently (<5%). These results are especially surprising in the case of type 2 layer 5 PTNs, which should have also been observed in that study. The discrepancies between our findings and those of (*6*) were probably related technical issues such as the low acquisition rate and the lack of adequate sparse labeling in (*6*).

The R/L hemi-tree compartmentalization is perhaps one of the most distinctive and intriguing properties of type 1 PTNs, which was not anticipated from previous work. This hemi-tree tuft compartmentalization enables PTNs to represent different sets of information in parallel, with each hemi-tree routing information to the soma independently acting as "a neuron within a neuron." It is conceivable that in larger and more complex primate and human PTNs, this property would have an even larger impact, with a greater number of almost isolated integrative zones in the tuft (*50*, *51*).

On the basis of our data, we propose a new integration and representation scheme of motor variables in tuft dendrites of M1 layer 5 PTNs. Motor variables are not represented in fully compartmentalized, small, nonoverlapping dendritic segments as previously reported (*6*). Instead, a given motor behavior or a sequence is represented by the activation of a specific combination of distal tuft segments, which are mutually coamplified via NMDA spikes to form spatial dendritic amplification maps for different motor behaviors. In this framework, the tuft tree of type 1 layer 5 PTNs is capable of dynamic combinatorial representation of a large number of motor variables and sequences within the same dendritic tuft branches (*2*, *52*, *53*).

REFERENCES AND NOTES

1. T. Branco, M. Häusser, *Curr. Opin. Neurobiol.* **20**, 494–502 (2010).
2. X. E. Wu, B. W. Mel, *Neuron* **62**, 31–41 (2009).
3. G. J. Stuart, N. Spruston, *Nat. Neurosci.* **18**, 1713–1721 (2015).
4. G. Major, M. E. Larkum, J. Schiller, *Annu. Rev. Neurosci.* **36**, 1–24 (2013).
5. S. D. Antic, W. L. Zhou, A. R. Moore, S. M. Short, K. D. Ikonomu, *J. Neurosci. Res.* **88**, 2991–3001 (2010).
6. J. Cichon, W. B. Gan, *Nature* **520**, 180–185 (2015).
7. L. M. Palmer *et al.*, *Nat. Neurosci.* **17**, 383–390 (2014).
8. L. Beaulieu-Laroche, E. H. S. Toloza, N. J. Brown, M. T. Harnett, *Neuron* **103**, 235–241.e4 (2019).
9. V. Francioni, Z. Padamsey, N. L. Rochefort, *eLife* **8**, e49145 (2019).
10. J. Voigts, M. T. Harnett, *Neuron* **105**, 237–245.e4 (2020).
11. A. Kerlin *et al.*, *eLife* **8**, e46966 (2019).
12. N. L. Xu *et al.*, *Nature* **492**, 247–251 (2012).
13. C. O. Lacefield, E. A. Pnevmatikakis, L. Paninski, R. M. Bruno, *Cell Rep.* **26**, 2000–2008.e2 (2019).
14. B. B. Ujfalussy, J. K. Makara, M. Lengyel, T. Branco, *Neuron* **100**, 579–592.e5 (2018).
15. R. Naud, B. Bathellier, W. Gerstner, *Front. Comput. Neurosci.* **8**, 90 (2014).
16. P. Poirazi, T. Brannon, B. W. Mel, *Neuron* **37**, 989–999 (2003).
17. D. Beniaguev, I. Segev, M. London, *Neuron* **109**, 2727–2739.e3 (2021).
18. A. Payeur, J. C. Béïque, R. Naud, *Curr. Opin. Neurobiol.* **58**, 78–85 (2019).
19. G. Kastellakis, P. Poirazi, *Front. Mol. Neurosci.* **12**, 300 (2019).
20. R. Legenstein, W. Maass, *J. Neurosci.* **31**, 10787–10802 (2011).
21. J. Bono, C. Clopath, *Nat. Commun.* **8**, 706 (2017).
22. B. W. Mel, in *Advances in Neural Information Processing Systems*, J. Moody, S. Hanson, R. Lippmann, Eds. (Morgan Kaufmann, 1992), vol. 4, pp. 35–42.
23. M. N. Economo *et al.*, *Nature* **563**, 79–84 (2018).
24. H. Markram *et al.*, *Cell* **163**, 456–492 (2015).
25. W. J. Gao, Z. H. Zheng, *J. Comp. Neurol.* **476**, 174–185 (2004).
26. Y. Wang, M. Ye, X. Kuang, Y. Li, S. Hu, *IBRO Rep.* **5**, 74–90 (2018).
27. S. Jiang *et al.*, *Sci. Rep.* **10**, 7916 (2020).
28. G. M. Shepherd, *Nat. Rev. Neurosci.* **14**, 278–291 (2013).
29. A. Groh *et al.*, *Cereb. Cortex* **20**, 826–836 (2010).
30. E. J. Kim, A. L. Juavinett, E. M. Kyubwa, M. W. Jacobs, E. M. Callaway, *Neuron* **88**, 1253–1267 (2015).
31. A. M. Hattox, S. B. Nelson, *J. Neurophysiol.* **98**, 3330–3340 (2007).
32. B. Tasic *et al.*, *Nature* **563**, 72–78 (2018).
33. L. N. Fletcher, S. R. Williams, *Neuron* **101**, 76–90.e4 (2019).
34. S. Levy *et al.*, *Neuron* **107**, 954–971.e9 (2020).
35. F. Fuhrmann *et al.*, *Neuron* **86**, 1253–1264 (2015).
36. T. W. Chen *et al.*, *Nature* **499**, 295–300 (2013).
37. T. Deneux *et al.*, *Nat. Commun.* **7**, 12190 (2016).
38. N. Mantel, *Cancer Res.* **27**, 209–220 (1967).
39. B. Manly, *The Statistics of Natural Selection* (Chapman & Hall, 1985).
40. P. W. Mielke, *Biometrics* **34**, 277 (1978).
41. B. Engelhard *et al.*, *Nature* **570**, 509–513 (2019).
42. P. Ramkumar, B. Dekleva, S. Cooler, L. Miller, K. Kording, *PLOS ONE* **11**, e0160851 (2016).
43. B. A. Sauerbrei *et al.*, *Nature* **577**, 386–391 (2020).
44. A. Nashef, O. Cohen, R. Harel, Z. Israel, Y. Prut, *Cell Rep.* **27**, 2608–2619.e4 (2019).
45. M. Lavzin, S. Rapoport, A. Polsky, L. Garion, J. Schiller, *Nature* **490**, 397–401 (2012).
46. M. E. Larkum, T. Nevian, M. Sandler, A. Polsky, J. Schiller, *Science* **325**, 756–760 (2009).
47. M. Larkum, *Trends Neurosci.* **36**, 141–151 (2013).
48. H. Jia, N. L. Rochefort, X. Chen, A. Konnerth, *Nature* **464**, 1307–1312 (2010).
49. M. E. Sheffield, D. A. Dombeck, *Nature* **517**, 200–204 (2015).
50. L. Beaulieu-Laroche *et al.*, *Cell* **175**, 643–651.e14 (2018).
51. H. Mohan *et al.*, *Cereb. Cortex* **25**, 4839–4853 (2015).
52. R. Urbanczik, W. Senn, *Neuron* **81**, 521–528 (2014).
53. J. Hawkins, S. Ahmad, *Front. Neural Circuits* **10**, 23 (2016).

ACKNOWLEDGMENTS

We thank S. Marom and B. Engelhard for helpful comments on the manuscript. We also thank S. Schwartz for advice in statistical tests and S. Gafniel for movie editing. **Funding:** This study was partially supported by the Israeli Science Foundation (J.S. and Y.S.), Prince funds (J.S. and Y.S.), Rappaport Foundation (J.S. and Y.S.), and Zuckerman Postdoctoral Fellowship (N.C.). **Author contributions:** Designed research: J.S., N.C., Y.S. Performed experiments: Y.O., N.C., S.A., M.A. Analyzed data: S.A., Y.O., J.S., H.B., O.B., Y.S. Writing – original draft: J.S. Writing – review and editing: J.S., Y.O., S.A., A.P.-P., Y.S. Performed simulations: A.P.-P., S.A., J.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Computer code can be found in: https://github.com/SchillersLab/Dynamic-compartmental-computations-in-tuft-dendrites-.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn1421
Materials and Methods
Figs. S1 to S14
References (*54*–*60*)
MDAR Reproducibility Checklist
Movies S1 to S10

7 November 2021; accepted 11 March 2022
10.1126/science.abn1421

REPORTS

ORGANIC CHEMISTRY

Allylic C–H amination cross-coupling furnishes tertiary amines by electrophilic metal catalysis

Siraj Z. Ali, Brenna G. Budaitis†, Devon F. A. Fontaine†, Andria L. Pace, Jacob A. Garwin, M. Christina White*

Intermolecular cross-coupling of terminal olefins with secondary amines to form complex tertiary amines—a common motif in pharmaceuticals—remains a major challenge in chemical synthesis. Basic amine nucleophiles in nondirected, electrophilic metal–catalyzed aminations tend to bind to and thereby inhibit metal catalysts. We reasoned that an autoregulatory mechanism coupling the release of amine nucleophiles with catalyst turnover could enable functionalization without inhibiting metal-mediated heterolytic carbon-hydrogen cleavage. Here, we report a palladium(II)-catalyzed allylic carbon-hydrogen amination cross-coupling using this strategy, featuring 48 cyclic and acyclic secondary amines (10 pharmaceutically relevant cores) and 34 terminal olefins (bearing electrophilic functionality) to furnish 81 tertiary allylic amines, including 12 drug compounds and 10 complex drug derivatives, with excellent regio- and stereoselectivity (>20:1 linear:branched, >20:1 *E*:*Z*).

Amines are one of the most prevalent structures found in small-molecule therapeutics, with an estimated 43% of drug candidates containing aliphatic amines, 60% of which are tertiary (*1*–*4*). Classic methods for their syntheses include nucleophilic substitution reactions of alkyl halides (*1*, *2*), reductive amination of the corresponding aldehydes or ketones (*5*), and Tsuji-Trost allylic aminations of *E*-allylic alcohols, acetates, or carbonates (*6*, *7*). Although simple tertiary amines can be made easily through these methods (e.g., cinnamaldehyde-derived allylic amines), more-complex amines require the synthesis of preoxidized, unstable reagents and more-tailored, substrate-dependent approaches (Fig. 1A, middle right). For linear, aliphatic amines—prominent structures in pharmaceuticals and agrochemicals—the prefunctionalized coupling partners (e.g., aldehydes) are often derived from terminal olefins (*8*). The addition of amines to terminal olefins provides a direct route to these important compounds that uses robust commercial building blocks (*9*–*12*). Advances in tandem radical-mediated olefin functionalization-elimination processes underscore the advantages of olefin-amine coupling (*13*, *14*). Metal-mediated allylic C–H amination provides an orthogonal methodology, where greater control of reactivity, selectivity, and functional group compatibility may be achieved.

Sulfoxide-oxazoline–palladium(II) [SOX·Pd$(OAc)_2$] allylic C–H aminations proceeding via π-allyl–Pd(SOX) intermediates demonstrated the capacity to promote allylic C–H to C–N cross-coupling between complex, unactivated terminal olefins and *N*-triflyl (Tf) primary alkylamines with high reactivity (fragment-coupling conditions) and high regio- and stereoselectivity (>20:1 linear:branched, >20:1 *E*:*Z*) to generate protected, secondary allylic amines (Fig. 1A, middle left) (*15*). Such metal-promoted intermolecular C–H aminations are generally limited to amine sources that have one or more electron-withdrawing groups covalently bound to nitrogen [e.g., Tf, toluenesulfonyl (Ts), and 2,2,2-trichloroethoxysulfonyl (Tces)] and require further synthetic manipulations to generate tertiary amines (*16*–*20*). At high concentrations, unprotected basic aliphatic amines bind to electrophilic metals and may reduce them and/or inhibit key steps in the catalytic cycle, such as intermolecular C–H activation (Fig. 1A, bottom left) (*21*–*23*). Under the current paradigm, an electron-withdrawing group cannot be covalently appended to a secondary amine without deactivating it as a nucleophile, rendering C–H amination to furnish complex tertiary amines an unsolved problem. Here, we disclose a general strategy to fragment-couple basic amines with terminal olefins in electrophilic metal–mediated C–H activation catalysis to furnish complex tertiary allylic amines with preparative reactivity, >20:1 linear-selectivity, *E*-selectivity, and orthogonal scope to current methods (Fig. 1A, bottom middle) (*24*).

Inspired by recent advances in cross-coupling and photoredox catalysis, where reactive intermediates are generated in low concentrations to exploit rate differences between unproductive and productive pathways (*25*–*27*), we sought to identify a mechanism for a slow release of amine nucleophiles under electrophilic metal–mediated C–H cleavage catalysis. The established mechanism for SOX·Pd$(OAc)_2$ allylic C–H amination involves electrophilic Pd(II)-mediated heterolytic C–H cleavage to furnish a cationic π-allyl–Pd(SOX) complex followed by functionalization and quinone-mediated Pd(0) reoxidation (vide infra). Supporting the hypothesis that low concentrations of free amine do not strongly inhibit C–H cleavage within this manifold, SOX·Pd$(OAc)_2$ amination with *N*-triflylamine proceeded in the presence of catalytic secondary amine (10%), whereas stoichiometric amounts halted catalysis (table S1). Lewis acids, such as boron trifluoride (BF_3), have been leveraged as transient protecting groups that mask amines during electrophilic metal–catalyzed reactions that occur at remote functionality; however, their capacity to modulate functionalization at the amine has not been explored (*28*–*30*). Although amine-BF_3 complexes are not able to undergo facile deprotonation (*31*), at elevated temperatures, they hydrolyze to furnish amine tetrafluoroborate (HBF_4) salts (*32*, *33*). We hypothesized that the hydroxyphenolate generated during quinone-mediated Pd(0) oxidation and/or the more basic tertiary amine products may act as bases, deprotonating amine-HBF_4 salts in situ to generate low concentrations of free amine nucleophiles that are regulated by catalyst loading (Fig. 1A, bottom right).

Amine-BF_3 complexes are stable to silica chromatography and air, providing an excellent way to purify and store secondary amines. We first examined (±)-MeO-SOX·Pd$(OAc)_2$ amine cross-coupling of 4-phenyl piperidine (**1**) as an amine-BF_3 complex with commercially available allylcyclohexane (**2**), an unactivated olefin, under fragment-coupling conditions (1 equivalent of each partner) and observed encouraging yields of the tertiary amine as a salt, which was isolated as amine **3** upon basic workup (Fig. 1B, entry 1). Addition of dibutyl phosphate (dbp), a Brønsted acid additive previously shown to increase reactivity in allylic functionalizations with sulfoxide–Pd$(OAc)_2$ catalysis (*34*), improved the yield of **3** to 74% (entry 2). Varying the acid loading from 0.25 to 0.5 equivalents increased reactivity (83% yield; entry 3); however, overalkylation was noted for some substrates and could be suppressed with lower acid loadings. In situ BF_3 complexation with **1** worked with equal efficiency (entry 4). Consistent with the reaction proceeding via an amine-HBF_4 salt, **1**·HBF_4 afforded comparable yields to **1**·BF_3 at lower acid loadings (entry 5). Exploration of alternative amine salts showed that although hydrochloride (HCl) salts were not effective, those with weakly coordinating counterions [trifluoroacetate (TFA), Ts, and dbp] afforded promising product yields, suggesting that, with further development, they may be used for this strategy (table S2). Expectedly, use of stoichiometric

Roger Adams Laboratory, Department of Chemistry, University of Illinois, Urbana, IL 61801, USA.
*Corresponding author. Email: mcwhite7@illinois.edu
†These authors contributed equally to this work.

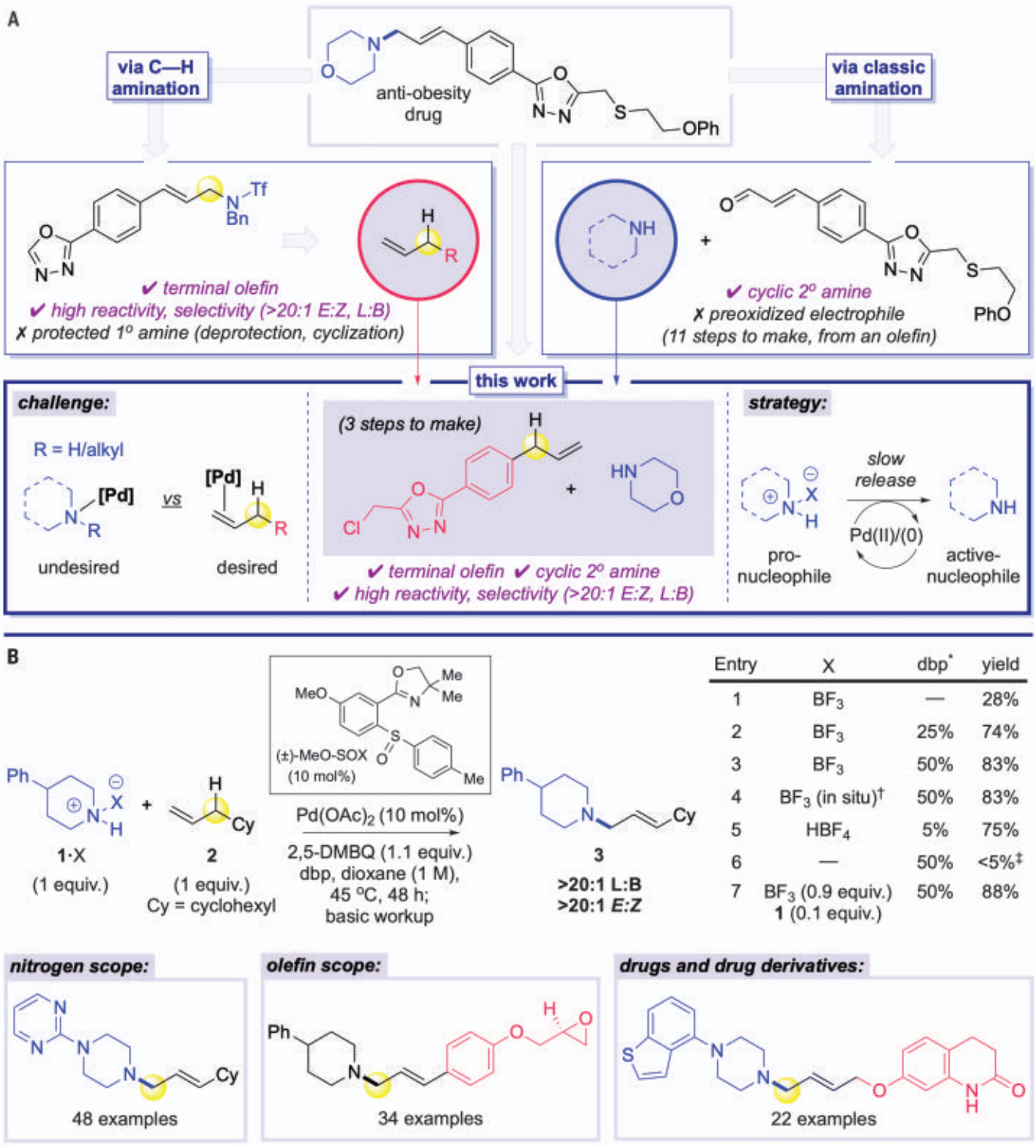


Fig. 1. Reaction design and development. (**A**) Merging the benefits of classic aminations (cyclic secondary amines) and C–H aminations (unactivated olefins) (middle) inspired the development of a fragment-coupling C–H amination to furnish complex tertiary amines (this work; bottom middle). The challenges associated with basic amines in electrophilic metal–mediated catalysis include undesired Pd(II) coordination that competes with C–H cleavage (bottom left). Our strategy uses catalyst turnover to regulate the generation of secondary amine nucleophiles (bottom right). (**B**) Reaction development and optimization studies using secondary amine-BF_3 complexes. All reactions were run on a 0.2-mmol scale open to air and moisture. All aminations proceeded in >20:1 linear:branched (L:B) and >20:1 *E*:*Z* selectivity, and free tertiary amines were isolated by means of a basic workup followed by column chromatography. Isolated yields are the average of two experiments. *Dibutyl phosphate. †The amine-BF_3 was complexed in the reaction vial without purification and was then subjected to AACC catalysis (supplementary materials). ‡Yield was determined by crude ^{1}H NMR analysis using benzotrifluoride as an internal standard.

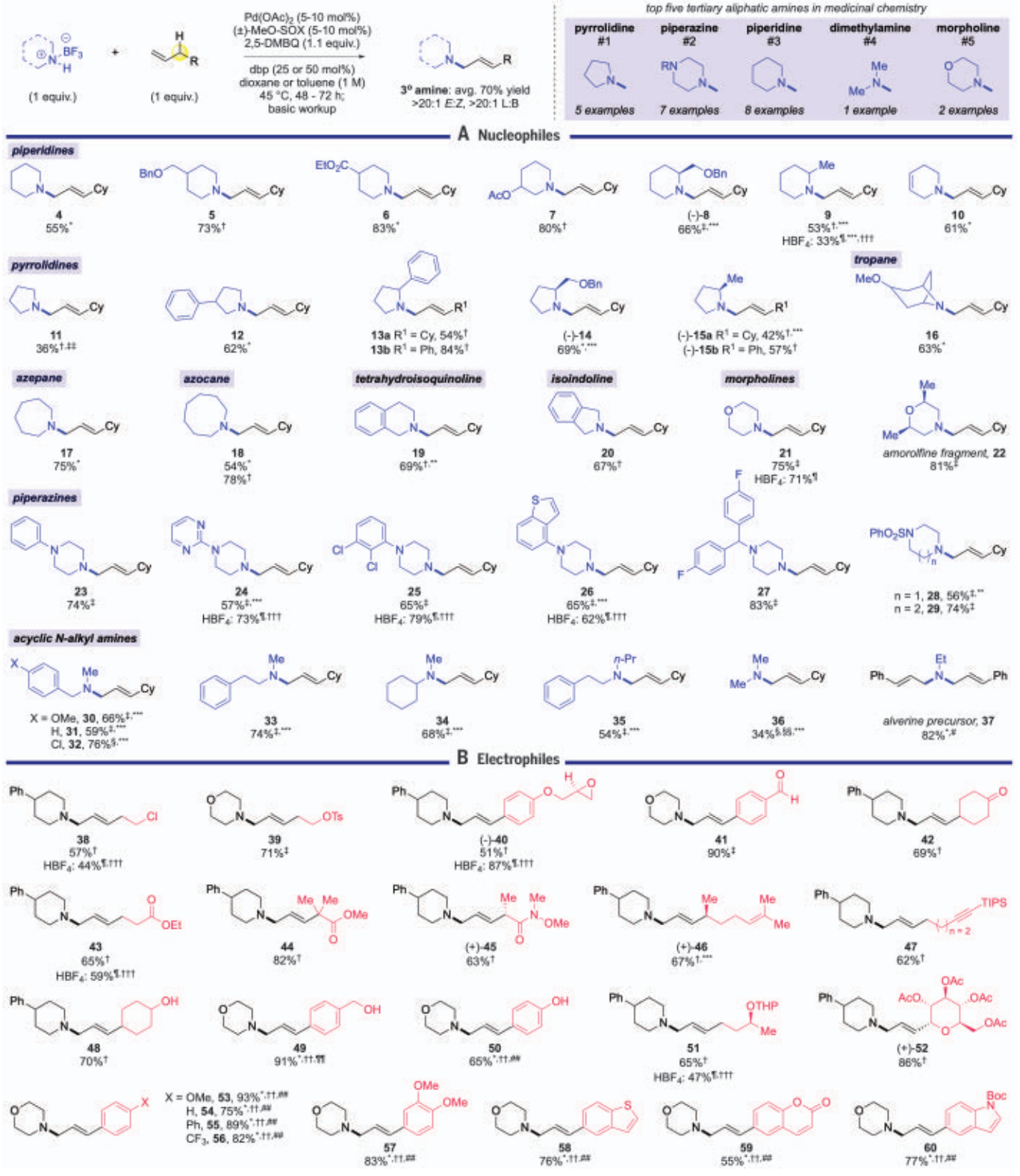


Fig. 2. AACC reaction scope. (**A**) Aliphatic cyclic and acyclic amine scope, including the top five aliphatic tertiary amines found in medicinal chemistry. (**B**) Terminal olefin scope. All reactions were run under ambient conditions. Unless otherwise indicated, all reactions were run on a 0.2-mmol scale using amine-BF_3 (1 equivalent), olefin (1 equivalent), $Pd(OAc)_2$ (10 mol %), (±)-MeO-SOX (10 mol %), 2,5-dimethyl-1,4-benzoquinone (2,5-DMBQ) (1.1 equivalents), and dbp in solvent (1 M) at 45°C for 48 hours, followed by basic workup. All amination products were formed in >20:1 L:B and >20:1 *E*:*Z* selectivity, and isolated yields of the free tertiary amine products are the average of three experiments. *25% dbp in dioxane. †50% dbp in dioxane. ‡25% dbp in toluene (in some cases, toluene diminishes diene by-product formation). §50% dbp in toluene. ¶5% dbp in dioxane. #0.5 equivalents of amine-BF_3. **0.4-mmol scale. ††5 mol % (±)-MeO-SOX·$Pd(OAc)_2$. ‡‡0.167 M. §§0.4 M. ¶¶12 hours. ##24 hours. ***72 hours. †††Average of two experiments.

amounts of free amine **1** under the standard conditions afforded no observed product, whereas substoichiometric amounts of **1** (0.1 equivalents) with **1**·BF_3 (0.9 equivalents) afforded product **3** in comparable yields (entries 6 and 7). In all cases examined for allylic C–H amination cross-coupling (AACC) (vide infra), use of 1 equivalent of each cross-coupling partner under conditions open to air and moisture furnished preparative yields of tertiary allylic amine products as pure regio- (>20:1 linear: branched) and stereoisomers (>20:1 *E*:*Z*). These results contrast with allylic aminations proceeding through tandem olefin functionalization-elimination pathways, which

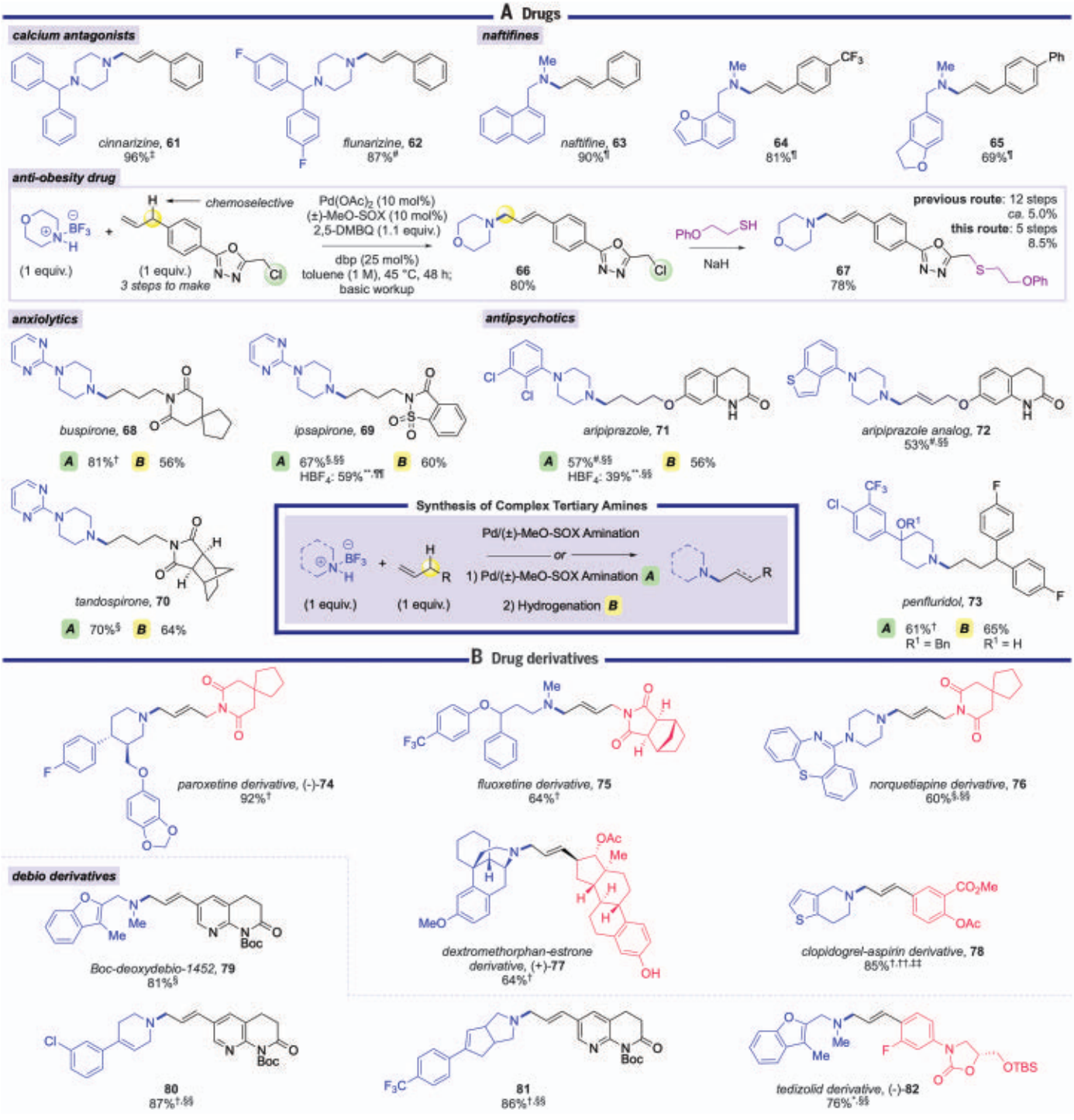


Fig. 3. Synthesis of complex tertiary aliphatic amine drugs and drug derivatives. (**A** and **B**) All reactions were run under ambient conditions. Unless otherwise indicated, all reactions were run on a 0.2-mmol scale using amine-BF_3 (1 equivalent), olefin (1 equivalent), $Pd(OAc)_2$ (10 mol %), (±)-MeO-SOX (10 mol %), 2,5-DMBQ (1.1 equivalents), and dbp in solvent (1 M) at 45°C for 48 hours, followed by basic workup. All amination products were formed in >20:1 L:B and >20:1 *E*:*Z* selectivity, and isolated yields of the free tertiary amine products are the average of three experiments. *25% dbp in dioxane. †50% dbp in dioxane. ‡25% dbp in toluene. §50% dbp in toluene. ¶50% dbp in methyl *tert*-butyl ether. #50% dbp in benzene. **5% dbp in dioxane. ††5 mol % (±)-MeO-SOX·$Pd(OAc)_2$. ‡‡24 hours. §§72 hours. ¶¶Average of two experiments.

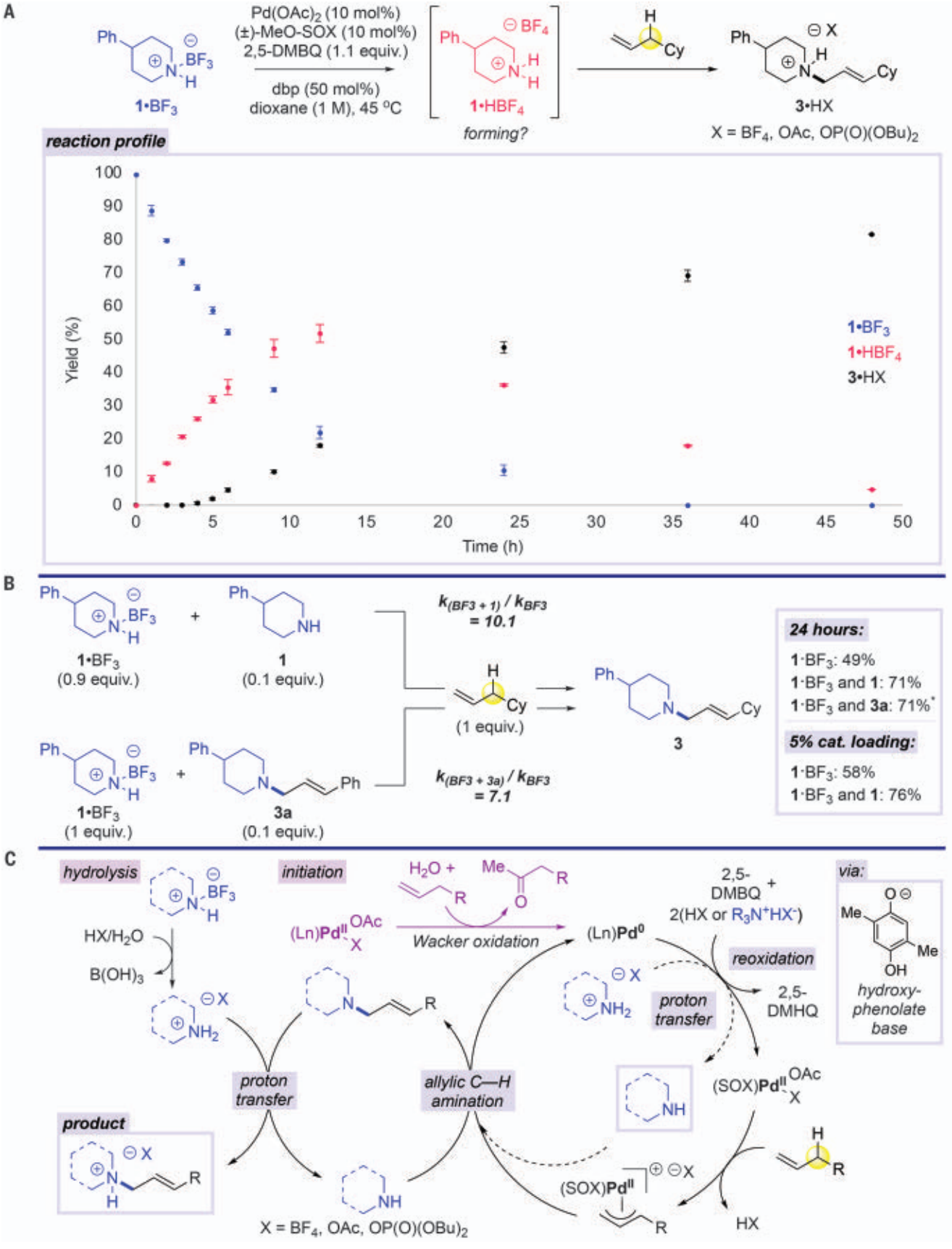


Fig. 4. Mechanistic studies and proposed mechanism. (**A**) Reaction profile monitored by quantitative ^{1}H NMR analysis using *(continued on next page)*

benzotrifluoride as an internal standard. All time points (excluding t = 0 and 48 hours) are the average of three experiments, with error bars indicating one standard deviation. The identity of **1**·HBF_4 was confirmed by matching its spectra to authentic material; however, we cannot exclude that other counterions (X) may be present. (**B**) Experiments to investigate the formation of free secondary amine nucleophile and the role of free tertiary amine in proton transfer as well as experiments to investigate reactivity at 5 mol % catalyst loading (cat. loading). See supplementary materials for details. *Yield was determined by crude ^{1}H NMR analysis using mesitylene as an internal standard. (**C**) Proposed mechanism for AACC. 2,5-DMHQ, 2,5-dimethylhydroquinone.

underscores the high reactivity and selectivity achieved with metal-mediated allylic C–H aminations.

Cyclic amines account for 59% of nitrogen moieties in pharmaceuticals (*3*). Five of the most common secondary amines found in aliphatic tertiary amine drug candidates (*1*) underwent AACC catalysis with unactivated terminal olefin **2** under cross-coupling stoichiometries (1 equivalent amine, 1 equivalent olefin) in preparative yields (Fig. 2A). Although unsubstituted pyrrolidine and piperidine nucleophiles showed modest amination yields as a result of overalkylation (**4** and **11**), substitution on the rings with varying functionality afforded increased yields (**5** to **10** and **12** to **15**). Ethers, esters, and acetates were all compatible at remote and/or proximal sites of these heterocycles, demonstrating that Lewis-basic oxygen functionality does not negatively affect amine-BF_3 complexation (**5** to **8** and **14**). α-Substituted piperidines and pyrrolidines generally afforded tertiary amines in preparative yields, demonstrating the tolerance of increased steric-bulk proximal to nitrogen (**8**, **9**, and **13** to **15**). Despite the greater steric bulk of phenyl versus methyl, 2-phenylpyrrolidine afforded tertiary amine products in higher yields than its 2-methyl counterpart with both unactivated and activated olefins (**13a** and **13b** versus **15a** and **15b**). Underscoring the high chemoselectivity of this allylic amination manifold, internal olefins were unreactive and could be incorporated on the nucleophile (**10**) as well as the olefin cross-coupling partner (vida infra; **46** and **47**). A topologically complex alkaloid core, tropane, as well as medium-sized heterocycles azepane and azocane were cross-coupled to olefin **2** in useful yields (**16** to **18**). Finally, cyclic amines containing low bond dissociation energy (BDE) benzylic α-amino C–H bonds, such as tetrahydroisoquinoline and isoindoline, were functionalized at nitrogen under this oxidative, two-electron process (**19** and **20**).

We next examined cyclic secondary amine nucleophiles bearing additional Lewis-basic heteroatom functionality, which introduces competing sites for BF_3 complexation or Pd(II) coordination. Ethereal oxygen, less basic than nitrogen, did not inhibit complexation: Morpholine-BF_3 as well as a disubstituted analog (core structure in antifungals amorolfine, fenpropimorph, and tridemorph) successfully underwent the reaction in preparative yields (**21** and **22**). Piperazines, the second most common tertiary aliphatic amine within medicinal chemistry, frequently contain aryl and/or alkyl substituents at the N1 and N4 positions (*1*, *3*, *4*). In *N*-aryl piperazines, the secondary aliphatic amine is preferentially complexed in the presence of a less basic tertiary aniline. Several *N*-aryl piperazines, including those substituted with pyrimidine, halogenated aromatics, and benzothiophene functionalities were generally effective nucleophiles as seen in their cross-coupling with unactivated terminal olefin **2** (**23** to **26**). In exploring *N*-alkyl piperazines, which contain basic tertiary aliphatic amines, diphenylmethyl piperazine underwent allylic functionalization to afford **27** in good yields, whereas less–sterically hindered benzyl piperazine afforded no cross-coupled product (supplementary materials). Piperazine nucleophiles with a strong inductively withdrawing group on one of the nitrogen atoms, such as phenylsulfonyl piperazine and its seven-membered ring homolog, afforded the desired products **28** and **29** in 56 and 74% yields, respectively. Amine-HBF_4 salts generally gave similar or diminished yields relative to their amine-BF_3 counterparts (**9**, **21**, and **26**; vide infra, **38**, **43**, **51**, **69**, and **71**); however, diene formation observed with **24** and **25** was suppressed under the HBF_4 manifold and led to increased yields (supplementary materials).

We subsequently investigated acyclic secondary amine nucleophiles that may face additional challenges because of their greater propensity for β-hydride elimination (*35*). *N*-methylbenzylamines, a common pharmacophore (*36*), readily underwent BF_3 complexation and afforded good yields of cross-coupled products irrespective of the electronic substitution of the benzyl moiety (**30** to **32**). Exchange of the benzyl moiety with phenethyl or sterically demanding cyclohexyl furnished tertiary amine products with analogous efficiencies (**33** and **34**). Replacement of the *N*-methyl moiety with *N*-propyl also furnished coupled product **35** in a useful yield. Dimethylamine—a small bulk commodity chemical and convenient precursor to dimethylamino groups prevalent in drug compounds—afforded the cross-coupled tertiary amine product **36** in useful 34% yield with the dialkylated, quaternary amine salt as a byproduct. Although primary amine-BF_3 nucleophiles yield dialkylation products under AACC catalysis, such reactivity may be used to furnish symmetrical tertiary amines: One equivalent of ethylamine-BF_3 was reacted with two equivalents of allylbenzene to afford a streamlined synthesis of the allylic precursor **37** to the smooth muscle relaxant alverine (*37*).

We next used a piperidine and morpholine, two of the most medicinally relevant amines, to explore the olefin scope for generality and orthogonal functional group compatibility relative to other methods for accessing tertiary amines (Fig. 2, top, and Fig. 2B) (*1*). There is ample precedent for basic secondary amines to displace leaving groups, such as halogens, sulfonates, and epoxide oxygens at C_{sp3} centers in Hofmann alkylations, to furnish tertiary amines (*1*, *2*). Under AACC catalysis, where the majority of the basic amine is deactivated as a salt, such functionality is maintained with chemoselective C–N bond formation occurring at the allylic terminus (**38**, **39**, and **40**). In a substrate bearing a terminal epoxide, the C–H allylic amination product **40** formed alongside dbp by-products from epoxide opening (supplementary materials). Notably, switching to the HBF_4 protocol at lower dbp loadings (5%) afforded **40** in 87% yield, despite the precedent for amine-HBF_4 salts to act as catalysts for epoxide-opening polymerization processes (*32*).

Remote carbonyl functionality, generally reactive with secondary amines and challenging to maintain in reductive or radical-based amination methods (*5*, *9*–*14*), was compatible with the oxidative conditions of AACC catalysis. Olefins containing aldehyde, ketone, ethyl ester, and gem-dimethyl ester functionalities afforded excellent yields of tertiary amine products (**41** to **44**). An olefin bearing Weinreb amide functionality with an α-stereocenter, an attractive synthetic handle, furnished chiral coupled product **45** that could be further elaborated.

Functionality prone to oxidation was also compatible with oxidative AACC catalysis. Substrates containing remote π-functionality, such as a citronellal-derived internal olefin and an internal alkyne, afforded the desired products **46** and **47** in synthetically useful yields with high chemoselectivity for the terminal olefin. Olefins bearing unprotected secondary and primary benzylic alcohols, known to undergo oxidation under Pd(II) catalysis (*38*), successfully furnished aminated products **48** and **49** in excellent yields with no detected carbonyl products. Phenol, a precedented nucleophile in Pd(0)-catalyzed allylic substitutions (*39*), was compatible, affording allylic amine **50** as the only observed product. Olefins with acid-labile functionality, such as a remote tetrahydropyranyl ether and a densely functionalized sugar,

were stable under the mildly acidic conditions of AACC catalysis to furnish **51** and **52** in preparative yields.

Allylbenzene olefins, which react via more-activated π-allyl–Pd(SOX) intermediates, underwent AACC catalysis at lower (±)-MeO-SOX·$Pd(OAc)_2$ loadings (5 mol %) and shorter reaction times (12 or 24 hours). Electron-rich, -neutral, and -poor allylbenzene derivatives uniformly furnished the allylic tertiary amine products in excellent yields (**53** to **57**). Common heterocycle motifs found in pharmaceuticals such as benzothiophene, coumarin, and indole—readily accessible as allylbenzenes but not as cinnamaldehydes—afforded allylic amine products in good yields (**58** to **60**). Collectively, the mild oxidative nature of this AACC catalysis provides an orthogonal approach to Hofmann alkylations, reductive aminations, olefin functionalizations, and allylic substitutions for synthesizing allylic tertiary amines.

We next evaluated the capacity of this allylic C–H amination cross-coupling to directly construct tertiary amine–bearing pharmaceuticals (Fig. 3A). Starting from commercial cyclizine fragments and allylbenzene, calcium antagonists cinnarizine (**61**) and flunarizine (**62**) were accessed in useful yields through AACC catalysis. Allylamine antifungal drugs naftifine (**63**) and known analogs (**64** and **65**) were furnished in high yields in the cross-coupling of readily accessible *N*-methylbenzylamines with allylated aromatics (*40*). Further showcasing the notable chemoselectivity for terminal olefins over traditional electrophiles, cross-coupling of morpholine with a terminal olefin bearing a reactive benzyl chloride electrophile afforded allylic amine **66** in 80% yield. Subsequent substitution to install an ethanethiol moiety afforded a streamlined synthesis of the experimental antiobesity compound **67** in approximately half the step count and twice the overall yield of the previous reductive amination route (*41*) (Fig. 3 versus Fig. 1A, middle right).

Linkage of a piperidine or piperazine to another heterocycle or aromatic moiety with a flexible 3- or 4-carbon chain is a characteristic feature of many psychiatric medicines (*42*). Buspirone (**68**), ipsapirone (**69**), and tandospirone (**70**), members of the anxiolytic drug class, were synthesized by cross-coupling pyrimidinylpiperazine-BF_3 to the corresponding alkylated imide olefin followed by hydrogenation. The syntheses of clinical antipsychotics aripiprazole (**71**) and its analog (**72**) were rapidly achieved with this approach, featuring notable functional group tolerance of an *O*-alkylated hydroxy-dihydroquinolinone electrophile (*43*, *44*). Penfluridol (**73**), a clinical diphenylbutylpiperidine antipsychotic, was additionally accessed through the AACC catalysis-hydrogenation sequence in useful yields. The amination proceeded smoothly with a piperidine nucleophile bearing an ionizable, tertiary benzyl–protected alcohol and provided a product prone to olefin isomerization.

Numerous drugs or their derivatives contain secondary amines that can be used in late-stage allylic C–H amination cross-coupling with allylated drug fragments to rapidly furnish complex tertiary amines in medicinally relevant settings (Fig. 3B). Serotonin reuptake inhibitors paroxetine, fluoxetine, and norquetiapine were readily cross-coupled with buspirone and tandospirone fragments via 4-carbon linkers to afford **74**, **75**, and **76** in useful yields. Cough suppressant dextromethorphan was coupled as its secondary amine to an allylated estrone derivative to generate **77** in 64% yield. The amine fragment of clopidogrel, a World Health Organization (WHO) essential drug, was cross-coupled with an allylated acetylsalicylic acid derivative to give **78**.

AACC catalysis is also well suited for the expedient generation of drug analogs. Debio-1452, an antibiotic in clinical trials (*45*), was readily synthesized as the Boc-deoxydebio-1452 analog (**79**) in the cross-coupling of *N*-methyl-benzofuranyl amine with allyl dihydronaphthyridinone. Tetrahydropyridine and pyrrolidine derivatives (**80** and **81**) were also accessed, further underscoring the high chemoselectivity for terminal versus internal olefins with this method. Alternatively, the debio-amine *N*-methyl-benzofuranyl fragment can readily be coupled to other olefin partners, including an allylated derivative of the broad-spectrum antibiotic tedizolid (**82**). These tertiary amine drugs and their derivatives were readily furnished through AACC catalysis using one catalyst and robust amine and olefin coupling partners under atmospheric conditions that are amenable to high-throughput, fragment-based drug discovery campaigns (*24*).

Mechanistic studies focused on determining how functionalization of the π-allyl–Pd [(±)-MeO-SOX] intermediate proceeded with amine-BF_3 pronucleophiles (Fig. 4). Amine-BF_3 complexes are hydrolyzed at elevated temperatures with ambient water to furnish amine-HBF_4 and amine–boric acid complexes (3:1 stoichiometry) (*32*, *33*). To investigate an amine-HBF_4 salt as an intermediate, we monitored the amount of amine-BF_3 (**1**·BF_3), amine-HBF_4 (**1**·HBF_4), and allylic C–H amination product (**3**·HX) in the reaction over time using quantitative proton nuclear magnetic resonance (^{1}H NMR) analysis (Fig. 4A). Initially, **1**·BF_3 was consumed steadily, with rapid formation of **1**·HBF_4. Only after an appreciable amount of **1**·HBF_4 was formed (~20 to 30%) did product **3**·HX formation occur. At the end of the reaction, tertiary amine product **3** was observed predominantly as an amine-HBF_4 salt that may serve to suppress deleterious overalkylation and amine-directed palladations of the product (*21*, *22*). We next sought to identify a mechanism whereby free amine nucleophile **1** could be generated from **1**·HBF_4 in the catalytic cycle. Gas chromatography–mass spectrometry (GC-MS) analysis early in the reaction showed trace methyl ketone formation, consistent with a Wacker olefin oxidation that reduces Pd(II) to Pd(0) (Fig. 4C) (supplementary materials) (*23*). Pd(0) oxidation by quinone provides a hydroxyphenolate base that may deprotonate **1**·HBF_4 to initiate amine functionalization. Consistent with this, a Pd(0) precatalyst may be used and furnishes product **3** in comparable yields (supplementary materials). After initiation, proton transfer from the **1**·HBF_4 to the more basic tertiary amine product **3** may be an additional slow-release mechanism for the nucleophile. Even though the free amine nucleophile **1** or tertiary amine product **3** were not directly detected, their involvement in catalysis is supported by 10- and sevenfold increases in initial rates, respectively (fig. S12), that manifested in increased yields at 24 hours when reactions were run with catalytic amounts of **1** or **3a** (10 mol %) (Fig. 4B). The increased rate afforded from catalytic secondary amine is beneficial in improving reaction efficiency: Preliminary investigations demonstrate the ability to reduce the catalyst loading (10 to 5 mol %) and maintain preparative yield. Although 10 mol % of free amine **1** was beneficial, at >20 mol % **1**, a substantial decrease in product yields was observed, likely as a result of catalyst inhibition (fig. S13). Key to AACC catalysis is the coupling of catalyst turnover to the generation of free amine that places an upper limit of ~10 to 20% free amine in the catalytic cycle at any given time. We anticipate that this general strategy may be broadly applicable to other electrophilic metal–mediated reactions using basic amine nucleophiles.

REFERENCES AND NOTES

1. S. D. Roughley, A. M. Jordan, *J. Med. Chem.* **54**, 3451–3479 (2011).
2. A. Trowbridge, S. M. Walton, M. J. Gaunt, *Chem. Rev.* **120**, 2613–2692 (2020).
3. E. Vitaku, D. T. Smith, J. T. Njardarson, *J. Med. Chem.* **57**, 10257–10274 (2014).
4. Y.-J. Wu, in *Progress in Heterocyclic Chemistry*, G. W. Gribble, J. A. Joule, Eds. (Elsevier, vol. 24, 2012), pp. 1–53.
5. E. Podyacheva, O. I. Afanasyev, A. A. Tsygankov, M. Makarova, D. Chusov, *Synthesis* **51**, 2667–2677 (2019).
6. I. D. G. Watson, A. K. Yudin, *J. Am. Chem. Soc.* **127**, 17516–17529 (2005).
7. J. Jing *et al*., *Chem. Commun.* **53**, 5151–5154 (2017).
8. A. Seayad *et al*., *Science* **297**, 1676–1678 (2002).
9. M. Utsunomiya, J. F. Hartwig, *J. Am. Chem. Soc.* **125**, 14286–14287 (2003).
10. S. Zhu, N. Niljianskul, S. L. Buchwald, *J. Am. Chem. Soc.* **135**, 15746–15749 (2013).
11. S. M. Bronner, R. H. Grubbs, *Chem. Sci.* **5**, 101–106 (2014).
12. A. J. Musacchio *et al*., *Science* **355**, 727–730 (2017).
13. Q. Cheng, J. Chen, S. Lin, T. Ritter, *J. Am. Chem. Soc.* **142**, 17287–17293 (2020).
14. D. J. Wang, K. Targos, Z. K. Wickens, *J. Am. Chem. Soc.* **143**, 21503–21510 (2021).
15. R. Ma, M. C. White, *J. Am. Chem. Soc.* **140**, 3202–3205 (2018).
16. T. Ide *et al*., *J. Am. Chem. Soc.* **143**, 14969–14975 (2021).
17. J. S. Burman, S. B. Blakey, *Angew. Chem. Int. Ed.* **56**, 13666–13669 (2017).
18. H. Lei, T. Rovis, *Nat. Chem.* **12**, 725–731 (2020).

19. T. Knecht, S. Mondal, J.-H. Ye, M. Das, F. Glorius, *Angew. Chem. Int. Ed.* **58**, 7117–7121 (2019).
20. E. T. Hennessy, R. Y. Liu, D. A. Iovan, R. A. Duncan, T. A. Betley, *Chem. Sci.* **5**, 1526–1532 (2014).
21. J. Rodrigalvarez *et al*., *Nat. Chem.* **12**, 76–81 (2020).
22. J. J. Topczewski, P. J. Cabrera, N. I. Saper, M. S. Sanford, *Nature* **531**, 220–224 (2016).
23. I. P. Beletskaya, A. V. Cheprakov, *Chem. Rev.* **100**, 3009–3066 (2000).
24. C. W. Murray, D. C. Rees, *Nat. Chem.* **1**, 187–192 (2009).
25. I. B. Perry *et al*., *Nature* **560**, 70–75 (2018).
26. D. M. Knapp, E. P. Gillis, M. D. Burke, *J. Am. Chem. Soc.* **131**, 6961–6963 (2009).
27. J. M. Dennis, N. A. White, R. Y. Liu, S. L. Buchwald, *ACS Catal.* **9**, 3822–3830 (2019).
28. J. M. Howell, K. Feng, J. R. Clark, L. J. Trzepkowski, M. C. White, *J. Am. Chem. Soc.* **137**, 14590–14593 (2015).
29. M. Lee, M. S. Sanford, *J. Am. Chem. Soc.* **137**, 12796–12799 (2015).
30. M. Ferrer, F. Sánchez-Baeza, A. Messeguer, A. Diez, M. Rubiralta, *J. Chem. Soc. Chem. Commun.* **1995**, 293–294 (1995).
31. J. F. Brown Jr., *J. Am. Chem. Soc.* **74**, 1219–1221 (1952).
32. R. E. Smith, C. H. Smith, *J. Appl. Polym. Sci.* **31**, 929–939 (1986).
33. P. Chabanne, L. Tighzert, J.-P. Pascault, B. Bonnetot, *J. Appl. Polym. Sci.* **49**, 685–699 (1993).
34. T. J. Osberger, M. C. White, *J. Am. Chem. Soc.* **136**, 11176–11181 (2014).
35. S. Wagaw, R. A. Rennels, S. L. Buchwald, *J. Am. Chem. Soc.* **119**, 8451–8458 (1997).
36. Y. He, H. Song, J. Chen, S. Zhu, *Nat. Commun.* **12**, 638 (2021).
37. R. N. Shakhmaev, A. Sh. Sunagatullina, V. V. Zorin, *Russ. J. Org. Chem.* **53**, 832–835 (2017).
38. D. R. Jensen, J. S. Pugsley, M. S. Sigman, *J. Am. Chem. Soc.* **123**, 7475–7576 (2001).
39. C. A. Discolo, A. G. Graves, D. R. Deardorff, *J. Org. Chem.* **82**, 1034–1045 (2017).
40. S. Ni *et al*., *J. Med. Chem.* **60**, 8145–8159 (2017).
41. J. Ammenn, J. R. Gillig, L. J. Heinz, P. A. Hipskind, M. D. Kinnick, Y.-S. Lai, J. M. Morin Jr., J. A. Nixon, C. Ott, K. A. Savin, T. Schotten, L. J. Slieker, N. J. Snyder, M. A. Robertson, "Multicyclic compounds for use as melanin concentrating hormone antagonists in the treatment of obesity and diabetes," International Patent WO 2003/097047 A1 (2003).
42. T. Heinrich *et al*., *J. Med. Chem.* **47**, 4684–4692 (2004).
43. H. Yamashita, J. Matsubara, K. Oshima, H. Kuroda, N. Ito, S. Miyamura, S. Shimizu, T. Tanaka, H. Takahashi, "Piperazine-substituted benzothiophenes for treatment of mental disorders," International Patent WO 2006/112464 A1 (2006).
44. T.-M. Tang *et al*., *Org. Chem. Front.* **8**, 3867–3875 (2021).
45. E. N. Parker *et al*., *Nat. Microbiol.* **5**, 67–75 (2020).

ACKNOWLEDGMENTS

We thank P. J. Hergenrother and E. Parker for helpful discussions on Debio-1452; S. E. Denmark for helpful discussions on the mechanism; L. Zhu for assistance with quantitative ^{1}H NMR analyses; the Fout laboratory for use of their GC-MS; and H. Shade for checking the procedure in Fig. 2, molecule **22**. **Funding:** Financial support for this work was provided by the NIH National Institute of General Medical Sciences (R35 GM122525). M.C.W. and S.Z.A. acknowledge the NIH (R01AI 136773-01) for support to evaluate Debio-1452 derivatives. The Bruker 500-Mz NMR spectrometer was obtained with the financial support of the Roy J. Carver Charitable Trust, Muscatine, IA, USA. **Author contributions:** S.Z.A. and M.C.W. conceived of the work. J.A.G. performed preliminary experiments. S.Z.A., B.G.B., D.F.A.F., A.L.P., and M.C.W. designed the experiments. S.Z.A., B.G.B., D.F.A.F., and A.L.P. performed and analyzed the experiments. S.Z.A., B.G.B., D.F.A.F., and M.C.W. wrote the manuscript. **Competing interests:** The University of Illinois has a patent (US 10,266,503 B1) on sulfoxide-oxazoline ligands for Pd(II)-catalyzed allylic C–H functionalizations. The authors declare no other competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Correspondence and requests for materials should be addressed to M.C.W. (mcwhite7@illinois.edu).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn8382
Materials and Methods
Figs. S1 to S13
Tables S1 and S2
NMR Spectra
References (*46*–*72*)

22 December 2021; accepted 11 March 2022
10.1126/science.abn8382

NATURAL HAZARDS

Citizen seismology helps decipher the 2021 Haiti earthquake

E. Calais[1,2,3,4]*, S. Symithe[3,5], T. Monfret[2,3,6], B. Delouis[2,3], A. Lomax[7], F. Courboulex[2,3], J. P. Ampuero[2,3], P. E. Lara[2,8], Q. Bletery[2,3], J. Chèze[2,3], F. Peix[2,3], A. Deschamps[2,3], B. de Lépinay[2,3], B. Raimbault[1], R. Jolivet[1,4], S. Paul[2,3,5], S. St Fleur[3,5], D. Boisson[3,5], Y. Fukushima[9], Z. Duputel[10], L. Xu[11], L. Meng[11]

On 14 August 2021, the moment magnitude (M_w) 7.2 Nippes earthquake in Haiti occurred within the same fault zone as its devastating 2010 M_w 7.0 predecessor, but struck the country when field access was limited by insecurity and conventional seismometers from the national network were inoperative. A network of citizen seismometers installed in 2019 provided near-field data critical to rapidly understand the mechanism of the mainshock and monitor its aftershock sequence. Their real-time data defined two aftershock clusters that coincide with two areas of coseismic slip derived from inversions of conventional seismological and geodetic data. Machine learning applied to data from the citizen seismometer closest to the mainshock allows us to forecast aftershocks as accurately as with the network-derived catalog. This shows the utility of citizen science contributing to our understanding of a major earthquake.

On 14 August 2021, a moment magnitude (M_w) 7.2 earthquake struck the southern peninsula of Haiti (Fig. 1A), leaving ~2500 people dead, 13,000 injured, at least 140,000 houses destroyed or damaged, and a number of water, sanitation, and health facilities severely affected (*1*). Because the earthquake affected an area that is mostly rural, with low population density, its impact was much lower than the smaller but devastating 12 January 2010 M_w 7.0 Haiti event (*2*–*4*). Most of the damage and casualties were concentrated in the populated cities of Les Cayes and Jérémie (Fig. 1B), but hard-to-reach rural communities also took a hit, in a context aggravated by the tropical storm that followed the event and chronic insecurity complicating field access from the capital city. In spite of these difficulties, and in the absence of an operational national network of conventional seismic stations, nearby seismological data were readily available during and after the earthquake because of a citizen seismology effort using inexpensive and low-maintenance "Raspberry Shake" (RS) seismic stations hosted by volunteers (*5*–*7*) (Fig. 1) [see (8), section 1]. This project had two original goals. The first was to install simple but scientifically useful seismological sensors in the homes of citizens to improve the dissemination of seismological information to the public, increase earthquake awareness, and promote grassroots protection initiatives (*8*). The second goal was to complement the national broadband seismological network, a high-technology system difficult to operate and maintain in a development context with a chronic lack of state resources. This citizen-based seismic network bears similarities to the Quake Catcher and Community Seismic networks deployed in California (*9*, *10*), although these use accelerometers only and are deployed in a region already well covered with conventional seismic stations. The 14 August 2021 earthquake and its aftershock sequence are an important test of the applicability of low-cost, citizen-hosted seismometers to provide scientifically relevant data for rapid response to a major earthquake.

The 2021 Nippes earthquake occurred within the Caribbean–North American plate boundary (Fig. 1A), where the two plates are converging obliquely at a speed of ~2 cm/year (*11*). The convergence component of plate motion is accommodated by the underthrusting of the North American oceanic lithosphere along the Puerto Rico Trench–North Hispaniola Fault, and the left-lateral component is accommodated by the Septentrional and Enriquillo strike-slip fault zones (*12*–*14*). The Enriquillo fault zone is considered the source of at least three major historical earthquakes occurring in 1701 [intensity magnitude (M_I) 6.6], 1751 (M_I 7.4), and 1770 (M_I 7.5) and a fourth, smaller earthquake in 1860 with M_I 6.3 (*15*, *16*) (Fig. 1A).

[1]Département de Géosciences, École Normale Supérieure, CNRS UMR 8538, PSL Université, Paris, France. [2]Université Côte d'Azur, Institut de Recherche pour le Développement, Centre National de la Recherche Scientifique, Observatoire de la Côte d'Azur, Géoazur, Valbonne, France. [3]CARIBACT Joint Research Laboratory, Université d'État d'Haïti, Université Côte d'Azur, Institut de Recherche pour le Développement, Port-au-Prince, Haïti. [4]Institut Universitaire de France, Paris, France. [5]URGéo, Faculté des Sciences, Université d'État d'Haïti, Port-au-Prince, Haïti. [6]Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar (ICM), CSIC, Barcelona, Spain. [7]ALomax Scientific, Mouans Sartoux, France. [8]Instituto Geofísico del Perú, Lima, Perú. [9]International Research Institute of Disaster Science, Tohoku University, Sendai, Japan. [10]Observatoire Volcanologique du Piton de la Fournaise, Université de Paris, Institut de Physique du Globe de Paris, CNRS, Paris, France. [11]Department of Earth, Planetary and Space Sciences, University of California, Los Angeles, CA, USA.
***Corresponding author. Email: eric.calais@ens.fr**

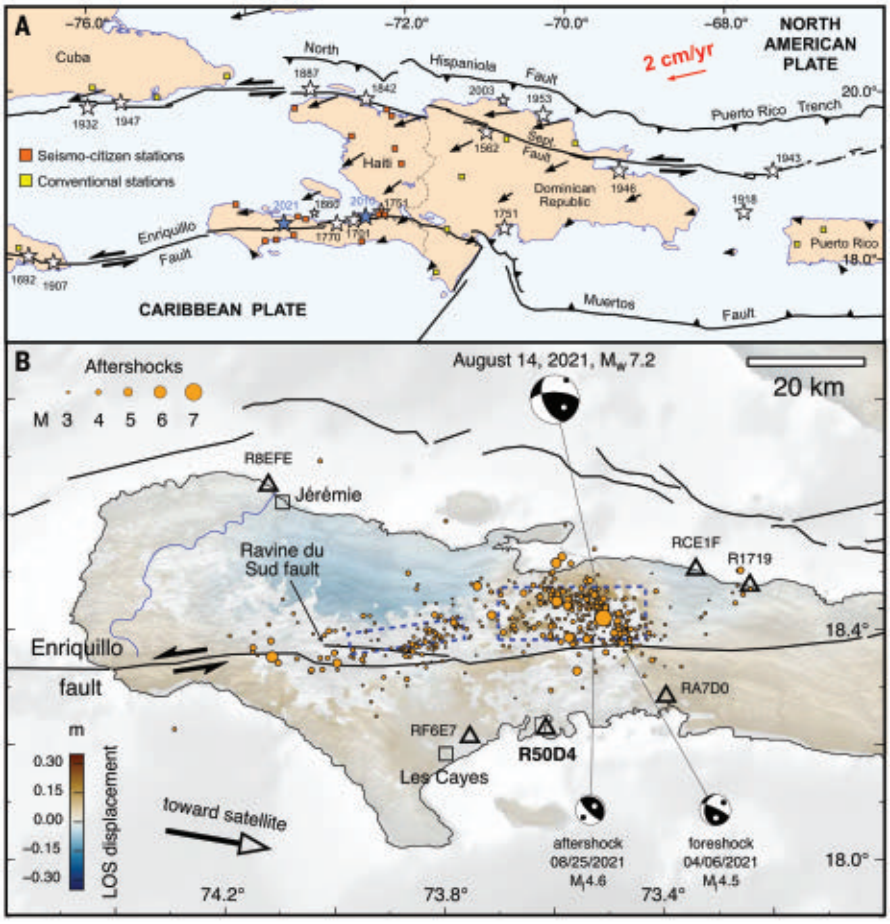


Fig. 1. Seismotectonic context of the 2021 Nippes earthquake. (A) Major active faults of the Caribbean–North America plate boundary zone with historical earthquakes (*16*) (stars) and Global Positioning System (GPS) velocities (black arrows) with respect to the Caribbean plate (*11*). (**B**) Relocated aftershock sequence (14 August to 9 September 2021) on top of a descending Sentinel interferogram spanning 3 to 8 August 2021. Triangles show citizen-hosted seismometers spanning the epicentral area. Line-of-sight (LOS) ground displacement north of the Enriquillo fault shows motion toward the satellite in the epicentral region (brown) and away from the satellite along the western part of the rupture (blue). Such reversal of the sense of motion along the LOS direction indicates substantial vertical motion in the epicentral region and almost pure horizontal, left-lateral motion to the west. Gray areas are not sufficiently coherent to ensure reliable phase unwrapping.

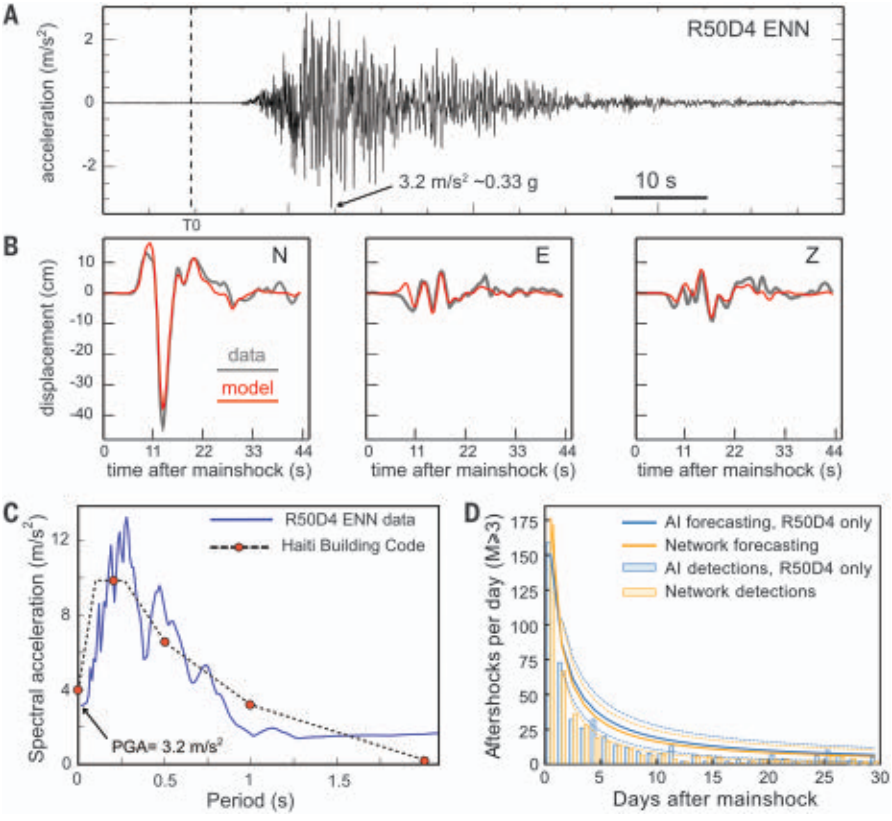


Fig. 2. Data and inferences from citizen station R50D4, 21 km from the 2021 Nippes earthquake rupture. (**A**) Signal in acceleration of the north component (channel ENN), which recorded a peak ground acceleration (PGA) of 0.33 g. Vertical line labeled T0 indicates the earthquake origin time. (**B**) Waveform fitting of the three components integrated to displacement and bandpass filtered between 0.06 and 0.5 Hz (N: north, E: east, Z: vertical up). The gray line is the observed signal, and the red line is the signal computed with the kinematic finite source model (Fig. 3B). (**C**) Spectral acceleration with 5% damping (blue line) of the north-south component of ground acceleration at the station (Fig. 1B). Red dots indicate the spectral values derived from the Haitian building code for the city of Les Cayes, closest to R50D4 and at the same distance from the rupture. The dashed line is drawn for visual interpretation but is not indicated in the code. Ground motion was stronger than expected for some frequency bands. (**D**) Detection and forecasting of aftershocks of magnitude >3 using the catalog derived from the whole network (orange) and from a single station (R50D4, blue). Histograms show detections, solid lines show forecast based on fitting an Omori-Utsu law to the first 12 hours of data, with their 95% confidence intervals indicated by dashed lines.

It was also the locus of the devastating M_w 7.0 earthquake of 12 January 2010. The epicentral region of the 2021 Nippes earthquake experienced two major events in 1952 (M_w 6.1) and 1953 (M_w 6.0) (*17*) and recurring clusters of smaller felt events, for example, the one in 2015 (*18*).

The mainshock of the 2021 Nippes earthquake was detected and characterized within minutes as M_w 7.2, consistent across most seismological agencies; this was 40% more energetic than the 2010 event, and with a source mechanism combining strike-slip and reverse faulting (*19*). It was recorded by five seismometers in Haiti: three RS stations hosted by citizens and two conventional stations in Port-au-Prince ~120 km from the epicenter, one US Geological Survey (USGS) accelerometer in the American embassy, and one educational broadband instrument in a high school (*20*). RS station R50D4, located 21 km from the epicenter (Fig. 1B), includes accelerometric sensors that recorded the mainshock without saturation with a maximum peak ground acceleration of 0.33 g on its north-south component (Fig. 2A). The high acceleration values for pseudo-periods lower than 0.5 s (Fig. 2C) [see (8), section 2] likely explains the severity of damages observed in the epicentral area in houses that, for the most part, were not built to earthquake-resistant standards. Spectral acceleration with 5% damping slightly exceeds the current Haiti building code (*21*, *22*) (Fig. 2C), indicating that even constructions built to current standards were exposed to an unexpectedly high hazard.

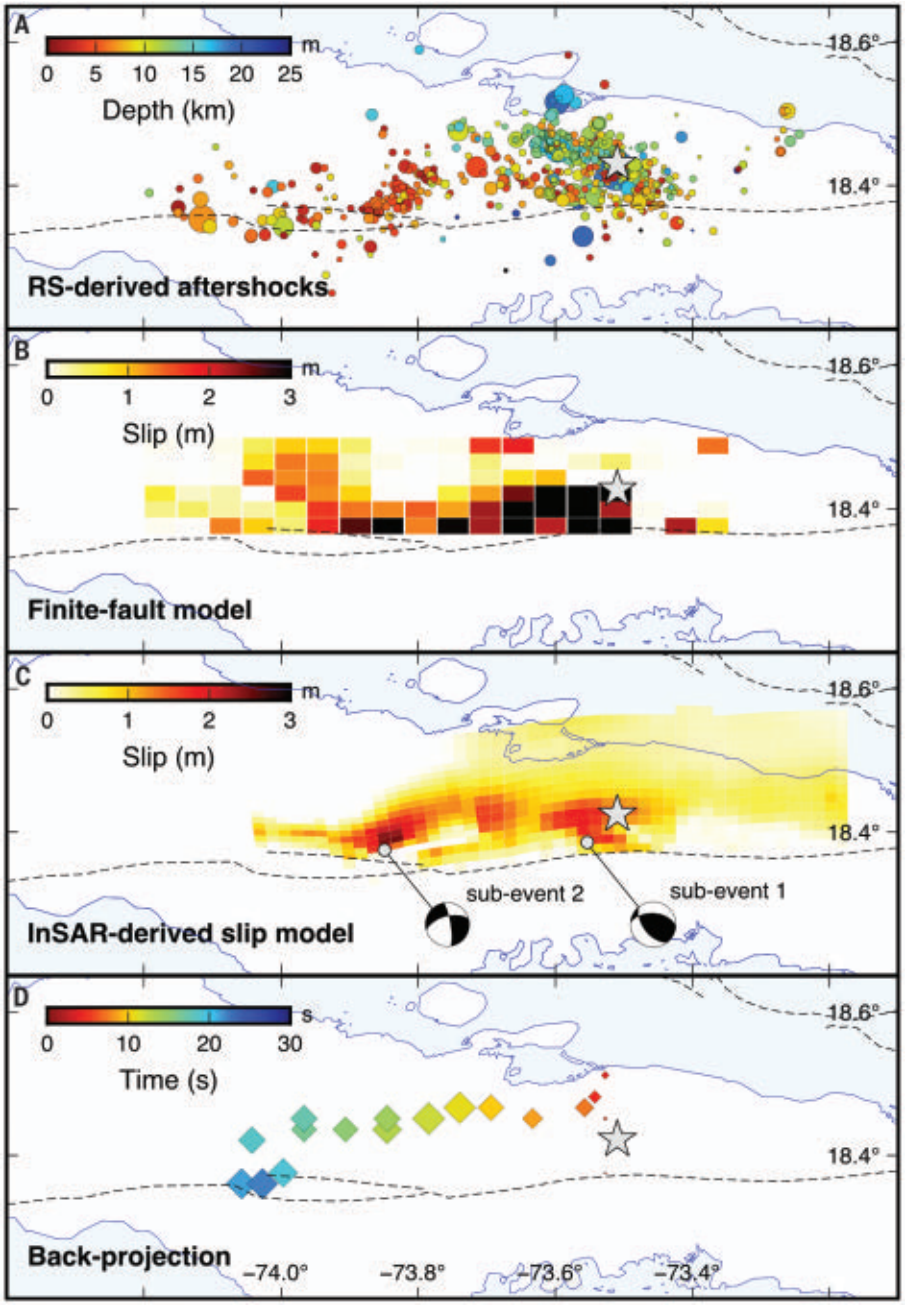


Fig. 3. Comparison between aftershock locations using citizen-hosted seismometers and the mainshock source mechanism. (**A**) Aftershock catalog after precise relocation with the 732 higher-quality events (14 August to 9 September 2021). (**B**) Kinematic finite fault model from an inversion of local and regional seismic stations. (**C**) Slip distribution inferred from InSAR data. The focal mechanisms derived from long-period modeling with two point sources are shown. (**D**) High-frequency (1 Hz) radiation sources (diamonds) from teleseismic back-projection source imaging. Symbol size is proportional to their relative energy and colored according to rupture time with respect to the mainshock. The gray star marks the 2021 Nippes epicenter from this study.

We determined a source mechanism for the mainshock using a linear finite-source model and the waveform inversion of data from conventional seismic stations at regional distance plus the near-source three-component accelerometric record from RS station R50D4 (Fig. 2B) [see (8), section 3]. The mechanism, consistent with global seismological agencies (*19*), combines 45% of strike-slip and 55% of reverse moment release, with an east-west trending nodal plane consistent with the local strike of the Enriquillo fault and dipping 60° to the north (Fig. 1B). The optimal centroid source depth was 6 km, indicating that most of the seismic moment was released at shallow depth. The citizen network detected two events of specific interest in the near vicinity of the mainshock. A possible foreshock on 6 April 2021, local magnitude (M_l) 4.5, coincides with the mainshock location, with a similar source mechanism (Fig. 1B). A substantial aftershock (08/25, M_l 4.6) detected by four RS stations is located within a few kilometers of the mainshock with a purely reverse mechanism (Fig. 1B). The three-component accelerometric recordings of the RS instruments were too noisy to be exploited at low frequency for these two smaller events, but their vertical velocimetric component contributed to the waveform inversion.

As of 9 September 2021, the citizen-based seismic network, together with regional conventional seismic stations located >120 km from the epicenter, detected 1031 aftershocks within a magnitude range of M_l 1.4 to 5.8, with a completeness magnitude around M_l 2.8. For comparison, 37 aftershocks are available for the same period in the global USGS catalog (*23*), which targets *M*4.5+ earthquakes only outside of the United States. We precisely relocated the mainshock and its aftershocks using manual (70% of events) and automatic

P-wave and S-wave arrival picks, source-specific station terms, and waveform similarity (*24*), with estimated error in absolute positions of 5 to 8 km and relative positions between nearby events of as little as 2 km [see (8), section 4]. We show the 732 higher-quality aftershock locations in Figs. 1B and 3A. We used only P-wave arrivals for the precise mainshock relocation because S-wave arrivals for large events are hidden in the P-wave train, and obtained the hypocenter at 18.42°N/73.51°W and 19 km depth.

Aftershocks are mostly located to the north of the Enriquillo fault (Figs. 1B and 3A), with the densest activity extending ~50 km east-west in two separate clusters: an eastern northwest-oriented cluster with ~4- to 20-km depth range, an ~10 × 25 km^2 area and overall dip to the north-northeast, containing the mainshock hypocenter at its base, and a western northeast-oriented cluster with an ~5 × 15 km^2 area and most events shallower than ~10 km depth. The western cluster merges westward into a sparse, east-west trend of events extending up to ~30 km along the Enriquillo fault zone, giving a total east-west extent of the main aftershock activity of as much as 80 km. Relocation without the citizen-based seismic network gives almost no depth constraint and produces a featureless cloud of epicenters of ~80 km extent and shifted ~20 km northeast of the centroid of the precisely located seismicity clusters.

The real-time detection of a large number of aftershocks permitted by the citizen-based seismic network allowed us to forecast their decay rates in a timely manner, information useful to the local population and emergency responders. The Reasenberg-Jones method (*25*) applied to the first 12 hours of the aftershock catalog shows a good match between the observed and forecast aftershock rates, which agree within 95% confidence over a 25-day interval [see (8), section 5]. In addition, we used a machine-learning (ML) approach to build an independent aftershock catalog using a single RS station (R50D4) [see (8), section 5]. These two independent catalogs are in good agreement, as well as the aftershock forecasts derived from each of them (Fig. 2D). This indicates that a single, well-located RS can provide the same forecast as the full network, maybe even a better one at very early times (fig. S6). This highlights the potential of low-cost instrumentation combined with ML for earthquake risk reduction in seismically active regions with limited resources.

We computed a kinematic finite fault-slip model using regional broad-band and strong-motion data, including near-field data from the R50D4 accelerometer (Fig. 3B) [see (8), section 6]. The rupture propagated unilaterally from the hypocenter westward over a distance of 50 to 60 km, at an average velocity of 2.8 km/s, with two areas of larger slip that correspond to the two aftershock clusters described above. The first area of large slip, to the east, is ~30 km long, with largely dominant reverse motion between 0 and 12 km depth. The second area of large slip, to the west, is limited to shallow depth (0 to 4 km) with pure left-lateral motion. The source time function indicates a rupture duration of ~20 s, followed by a small, separated, and less well-constrained burst near the western termination of the rupture. Teleseismic back-projection source imaging [see (8), section 7] yields first-order rupture characteristics consistent with the kinematic source inversion results, with a 50- to 60-km-long rupture propagating unilaterally westward at an average speed of ~3 km/s (Fig. 3D). This consistency relies on calibrating seismic ray propagation paths using aftershock data to account for local structure heterogeneity. The accuracy of the aftershock locations provided by citizen-based seismic stations was essential to ensuring the quality of the calibration.

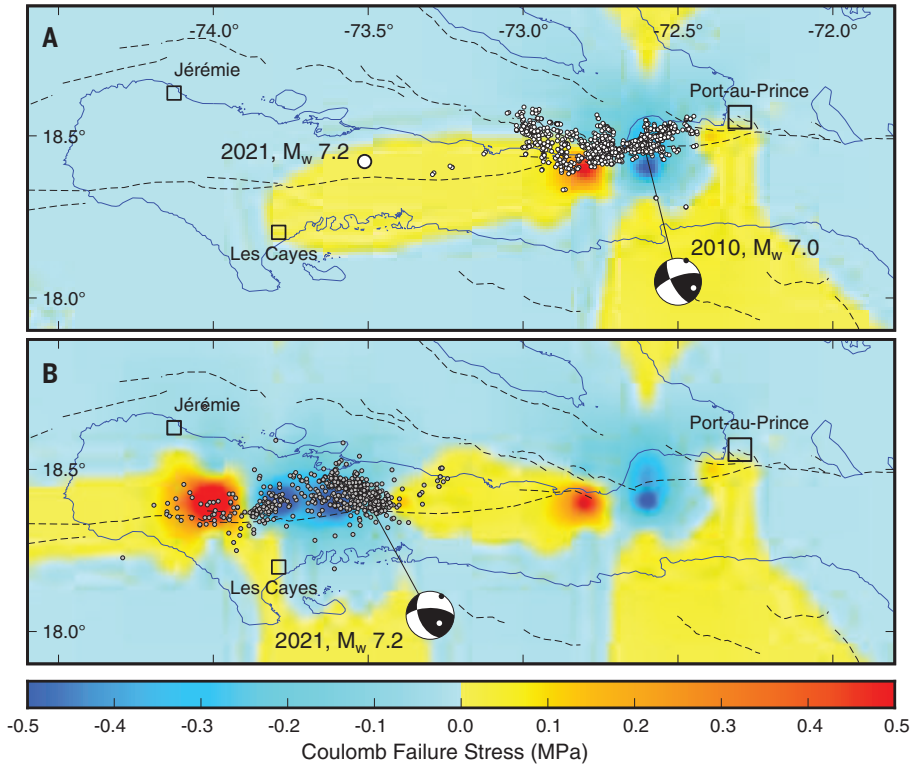


Fig. 4. CFS on east-west trending, vertical strike-slip faults. (**A**) CFS imparted by the 2010 earthquake, with its aftershocks shown as white dots. (**B**) CFS imparted by both the 2010 and 2021 earthquakes. The gray circles show the 2021 aftershock sequence as of 9 September 2021. The CFS is calculated at 5-km depth with a friction coefficient of 0.2.

We confirmed the seismic source mechanism using independent geodetic data available with a few weeks' delay [see (8), section 8]. Radar interferograms from the Sentinel 1 A and B and ALOS-2 satellites show substantial vertical motion in the epicentral area, consistent with thrusting on a north-dipping structure (Fig. 1B), and a rupture that reached the surface along the previously mapped Ravine du Sud fault (*26*) (Fig. 1B) but remained blind otherwise. A nonlinear least-squares search for the rupture geometry considering two rectangular fault planes [see (8), section 9] found that best-fit planes that coincide with the two aftershock clusters described above (Figs. 1B and 3A). A north-dipping (~60° north) plane in the eastern part of the epicentral region shows a combination of reverse and strike-slip motion, with a surface trace that coincides with the Enriquillo fault. A steeper (~71° north) north-dipping plane to the west shows mostly strike-slip motion, with a surface trace that coincides with the Ravine du Sud fault.

Given the coincidence between the nonlinear inversion rupture and the surface expression of the Enriquillo and Ravine du Sud faults, we used their mapped traces to build north-dipping rupture geometries at depth and infer the distribution of coseismic slip along them (Fig. 3C) [see (8), section 10]. The resulting interferometric synthetic aperture radar (InSAR) slip distribution is consistent with the rupture of two main patches, coinciding with the relocated aftershocks (Fig. 3A)

and with the finite fault seismic model (Fig. 3B). This consistency highlights the value of RS data to rapidly assess the main characteristics of this earthquake sequence.

In the slip models, the main patch to the east coincides with the mainshock epicenter location, with slip reaching 1.9 m, dominated by reverse motion. A second patch to the east coincides with the Ravine du Sud fault, with up to 2.3 m of purely strike-slip motion. The focal mechanisms corresponding to these two slip patches, highlighted by the aftershock distribution, are within uncertainties of those estimated independently from long-period modeling considering two point sources (Fig. 3C) [see (8), section 11].

We used this coseismic slip model, together with that of the 2010 earthquake (*27*, *28*), to compute the Coulomb failure stress (CFS; Fig. 4) imparted on faults of similar orientation and kinematics, as the main, strike-slip Enriquillo fault [see (8), section 12]. The initiation area of the 2021 rupture falls within an area of increased CFS caused by the 2010 event, an indication that the two earthquakes may be part of a sequence in which the 2010 event triggered the 2021 earthquake, as observed on other major strike-slip fault systems. The aftershock distribution of the two earthquakes shows that their ruptures are not contiguous. The ~60-km-long fault segment between them, as well as other segments to the west and east, have not ruptured in a major earthquake since at least the series of four events in the 18th century (*16*), and show increased CFS (Fig. 4). The 2010 and 2021 events have therefore increased earthquake hazard in southern Haiti, information critical to long-term planning for the region.

The 2021 Nippes earthquake bears similarities to the 2010 event (*2*, *3*, *27*, *28*). Both earthquakes exhibited aftershocks and coseismic slip north of the Enriquillo fault, initiated with a substantial component of reverse faulting motion on an eastern segment, and propagated westward with later, mostly strike-slip motion. Their marked dip-slip moment release is intriguing given the mainly strike-slip motion recorded geologically on the Enriquillo fault, information hard-wired into Haiti's seismic hazard map (*21*). It is consistent, however, with interseismic geodetic measurements (*11*, *29*, *30*) (Fig. 1B) and onshore and offshore geophysical data (*31*–*33*) showing far-field kinematics combining strike-slip and convergence, with north-northeast/south-southwest–directed compression. A reappraisal of the seismic hazard map of Haiti is therefore needed to account for this substantial north-south shortening component and to provide updated information for building code purposes.

The rapid assessment of the source mechanism, near-field ground shaking, and aftershock distribution of the 2021 Nippes earthquake was made possible by inexpensive seismometers hosted by citizens, together with information from classic seismological and geodetic data and models. The inclusion of the RS data in waveform inversions shows that they provide data of sufficient quality for adding valuable near-source information into the slip model, as confirmed by the InSAR slip inversion. This is an important example of a direct contribution of citizen seismology to understanding a large and damaging earthquake in the absence of conventional seismic stations in the near field of the event, highlighting the added value of citizen seismology for rapid earthquake response. The high benefit-to-cost ratio of citizen seismology makes it particularly relevant to regions of similar socioeconomic level as Haiti, where the implementation of conventional seismic networks operated by official institutions may be difficult (*34*).

REFERENCES AND NOTES

1. UN Office for the Coordination of Humanitarian Affairs, "Haïti: Tremblement de terre Rapport de Situation No. 2 Au 26 Août 2021" (OCHA, 2021); https://reliefweb.int/report/haiti/ha-ti-tremblement-de-terre-rapport-de-situation-no-2-au-26-ao-t-2021.
2. E. Calais *et al*., *Nat. Geosci.* **3**, 794–799 (2010).
3. G. P. Hayes *et al*., *Nat. Geosci.* **3**, 800–805 (2010).
4. M. Hashimoto, Y. Fukushima, Y. Fukahata, *Nat. Geosci.* **4**, 255–259 (2011).
5. E. Calais *et al*., *Front. Earth Sci. (Lausanne)* **8**, 542654 (2020).
6. R. E. Anthony, A. T. Ringler, D. C. Wilson, E. Wolin, *Seismol. Res. Lett.* **90**, 219–228 (2018).
7. Ayiti-Séismes Project, "Prognosis on 08/23/2021 of the aftershocks of the Nippes earthquake, Haiti (08/14/2021, magnitude 7.2)" (Ayiti-Séismes Project, 2021); https://ayiti.unice.fr/ayiti-seismes/
8. L. Fallou, E. Calais, A. Corbet, L. Hurbon, J. M. Théodat, "Citizen-seismology in Haiti, understanding citizens' interest and beliefs to enhance community resilience and contribute to risk reduction," paper presented at the Citizen Science SDG Conference: Knowledge for Change: A Decade of Citizen Science (2020-2030) in Support of the Sustainable Development Goals, Berlin, 14–15 October 2020.
9. E. S. Cochran, J. F. Lawrence, C. M. Christensen, R. S. Jakka, *Seismol. Res. Lett.* **80**, 26–30 (2009).
10. R. W. Clayton *et al*., *Ann. Geophys.* **54**, 6 (2011).
11. S. Symithe, E. Calais, J. B. de Chabalier, R. Robertson, M. Higgins, *J. Geophys. Res. Solid Earth* **120**, 120 (2015).
12. P. Mann, F. W. Taylor, R. L. Edwards, T. L. Ku, *Tectonophysics* **246**, 1–69 (1995).
13. E. Calais *et al*., *Geophys. Res. Lett.* **29**, 1856 (2002).
14. P. Mann *et al*., *Tectonics* **21**, 7–26 (2002).
15. J. Scherer, *Bull. Seismol. Soc. Am.* **2**, 161–180 (1912).
16. W. H. Bakun, C. H. Flores, U. S. ten Brink, *Bull. Seismol. Soc. Am.* **102**, 18–30 (2012).
17. I. Bondár, E. R. Engdahl, A. Villaseñor, J. Harris, D. Storchak, *Phys. Earth Planet. Inter.* **239**, 2–13 (2015).
18. C. Prépetit, "Anse-à-Veau, la ville sismique oubliée" (Bureau of Mines and Energy, Haiti, 2016); http://www.bme.gouv.ht/uts/Anse-à-Veau.pdf.
19. European-Mediterranean Seismological Centre, "M 7.2 - HAITI REGION - 2021-08-14 12:29:09 UTC" (EMSC, 2021); https://www.emsc-csem.org/Earthquake/earthquake.php?id=1023410#map.
20. F. Courboulex *et al*., *Seismol. Res. Lett.* **83**, 870–873 (2012).
21. A. Frankel, S. Harmsen, C. Mueller, E. Calais, J. Haase, *Earthq. Spectra* **27** (1_suppl1), S23–S41 (2011).
22. Ministère des Travaux Publics, Transports et Communications, "Code National du Bâtiment d'Haït (CNBH) 2012" (MTPTC, 2013); https://www.mtptc.gouv.ht/media/upload/doc/publications/CNBH_fusion.pdf
23. US Geological Survey, "Search earthquake catalog" (USGS, 2022); https://earthquake.usgs.gov/earthquakes/search/
24. A. Lomax, A. Savvaidis, *J. Geophys. Res. Solid Earth* **127**, e2021JB023190 (2022).
25. P. A. Reasenberg, L. M. Jones, *Science* **243**, 1173–1176 (1989).
26. N. Saint Fleur, N. Feuillet, Y. Klinger, *Tectonophysics* **771**, 228235 (2019).
27. S. J. Symithe, E. Calais, J. S. Haase, A. M. Freed, R. Douilly, *Bull. Seismol. Soc. Am.* **103**, 2326–2343 (2013).
28. R. Douilly *et al*., *Bull. Seismol. Soc. Am.* **103**, 2305–2325 (2013).
29. B. Benford, C. DeMets, E. Calais, *Geophys. J. Int.* **191**, 481–490 (2012).
30. S. Symithe, E. Calais, *Tectonophysics* **679**, 117–124 (2016).
31. J. Rodriguez, J. Havskov, M. B. Sørensen, L. F. Santos, *J. Seismol.* **22**, 883–896 (2018).
32. D. Possee *et al*., *Tectonics* **38**, 1138–1155 (2019).
33. J. Corbeau *et al*., *Tectonics* **35**, 1032–1046 (2016).
34. S. Subedi, G. Hetényi, P. Denton, A. Sauron, *Front. Earth Sci. (Lausanne)* **8**, 73 (2020).

ACKNOWLEDGMENTS

This project benefits from the collaboration of citizen seismologists in Haiti and the extra help of seismologists from Géoazur for manual picking of the aftershocks. J. Haase provided comments that improved an early version of the manuscript. ALOS-2 data were provided from JAXA through the Earthquake Working Group coordinated by the Geospatial Information Authority of Japan and JAXA. We acknowledge seismic data from regional networks in the Dominican Republic, Cuba, Jamaica, and Alaska, and thank their operating agencies for making them available. **Funding:** This work was supported by the Centre National de la Recherche Scientifique (CNRS) and the Institut de Recherche pour le Développement (IRD) through their "Natural Hazard" program (E.C., S.S., T.M., B.D., F.C., J.P.A., J.C., A.D., D.B., S.P.); the FEDER European Community program within the Interreg Caraïbes "PREST" project (E.C., S.S., D.B.); Institut Universitaire de France (E.C., R.J.); Université Côte d'Azur and the French Embassy in Haiti (S.P.); the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant no. 758210, Geo4D project to R.J. and grant no. 805256 to Z.D.); the French National Research Agency (project ANR-21-CE03-0010 "OSMOSE" to E.C. and ANR-15-IDEX-01 "UCAJEDI Investments in the Future" to Q.B.); the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant no. 949221 to Q.B.); and HPC resources of IDRIS (under allocations 2020-AD011012142, 2021-AP011012536, and 2021-A0101012314 to Q.B.). **Author contributions:** E.C. designed and coordinated the study. S.S., S.S.F., and D.B. collected the citizen-seismology data. S.P., F.C., T.M., A.D., V.C., J.C., and F.P. collected the mainshock and aftershock bulletins. A.L. relocated the earthquake catalog. P.L. and Q.B. performed the ML-based aftershock detections. B.D. prepared the point source, linear, and kinematic models. J.P.A. analyzed the aftershock forecast. F.C. performed the RS50D spectral analysis. L.X. and L.M. prepared the source back-projection. R.J. and B.R. performed the InSAR analysis and resulting fault model. Z.D. performed the multiple point source solution. Y.F. processed the ALOS-2 interferograms. All authors wrote the original version of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data and code used in this study are openly available. RADAR data can be obtained through ESA (Sentinel) or JAXA (Alos-2). Aftershock data can be obtained from https://ayiti.unice.fr/ayiti-seismes/ (*7*). The codes used to process or model the data are published and public (*8*). The catalog of high-precision earthquake relocated with the NLL-SSST-coherence procedure (SM4) is available as supplementary data.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn1045
Materials and Methods
Supplementary Text
Figs. S1 to S26
References (*35*–*99*)
Data S1

3 November 2021; accepted 23 February 2022
Published online 10 March 2022
10.1126/science.abn1045

CATALYSIS

Ambient-pressure synthesis of ethylene glycol catalyzed by C_{60}-buffered Cu/SiO_2

Jianwei Zheng[1]†, Lele Huang[1]†, Cun-Hao Cui[1], Zuo-Chang Chen[1], Xu-Feng Liu[1], Xinping Duan[1], Xin-Yi Cao[2], Tong-Zong Yang[3], Hongping Zhu[1], Kang Shi[1], Peng Du[1], Si-Wei Ying[1], Chang-Feng Zhu[3], Yuan-Gen Yao[2], Guo-Cong Guo[2], Youzhu Yuan[1]*, Su-Yuan Xie[1]*, Lan-Sun Zheng[1]

Bulk chemicals such as ethylene glycol (EG) can be industrially synthesized from either ethylene or syngas, but the latter undergoes a bottleneck reaction and requires high hydrogen pressures. We show that fullerene (exemplified by C_{60}) can act as an electron buffer for a copper-silica catalyst (Cu/SiO_2). Hydrogenation of dimethyl oxalate over a C_{60}-Cu/SiO_2 catalyst at ambient pressure and temperatures of 180° to 190°C had an EG yield of up to 98 ± 1%. In a kilogram-scale reaction, no deactivation of the catalyst was seen after 1000 hours. This mild route for the final step toward EG can be combined with the already-industrialized ambient reaction from syngas to the intermediate of dimethyl oxalate.

Ethylene glycol (EG) is commonly used as antifreeze and feedstock for polyethylene terephthalate used in bottles and packaging (*1*). In contrast to the production of petroleum-derived EG (*2*), EG can also be produced from syngas (CO and, more recently, CO_2 mixed with H_2) (*3*, *4*). Direct hydrogenation of syngas toward EG requires high pressure (>100 bar) at 230°C but has a low yield (theoretically 57% and experimentally <7%) resulting from thermodynamic limitation of Gibbs free energy and side reactions (*5*, *6*). Two-step strategies starting from syngas through intermediate species such as dimethyl oxalate (DMO) are suitable for EG production (Fig. 1A and fig. S1) (*7*–*10*). The coupling of CO with methanol to make DMO has been safely industrialized with a Pd catalyst approaching its thermodynamic yield limit at ambient pressure (*11*–*13*). However, conversion of DMO to EG still requires elevated hydrogen (H_2) pressure (typically 20 to 30 bar) and temperature (200°C). High-pressure reactions with H_2 compressors present environmental and safety risks that could be addressed with efficient catalysts running at near-ambient pressure conditions (*4*).

A copper-chromium catalyst was originally applied in the DMO-to-EG reaction but was detrimental to human health and the environment, primarily owing to the toxicity of chromium. In addition, high H_2 pressure (30 bar) was required (*14*). Alternatively, a copper-silica (Cu/SiO_2) catalyst that selectively hydrogenates the DMO-to-EG reaction may be used (*15*–*18*); however, Cu/SiO_2 suffers from insufficient activity at low pressure as well as poor stability (*7*, *9*, *14*, *17*). Efforts have been primarily directed toward promoting the Cu/SiO_2 catalyst by adding promoters such as B, Zn, and Au to tailor the electronic density of Cu species or enhance metal-support interactions (*16*, *19*–*22*). However, these element-modified catalysts readily deteriorated in structure and did not perform well at low H_2 pressure (*7*, *23*). These kinds of elemental promoters increase the density of electron-deficient copper by irreversibly accepting electrons from copper. In contrast to elemental promoters, molecular promoters such as C_{60} and C_{70}—in addition to accepting electrons from copper—can also give electron feedback to the electron-deficient copper to render useful redox properties during the catalytic process. Electronic modification by C_{60} is practical in photovoltaics, but the buffering effect of C_{60} is rarely reported in catalysis (*24*, *25*).

We used C_{60} as an electronic buffer to balance the electronic density of active Cu species and overcome the restrictions imposed on conventional Cu catalysts. As shown in Fig. 1, B and C, the conventional Cu/SiO_2 catalyst without C_{60} had a low EG yield of 9.6%, but DMO could be almost completely converted to EG with a yield of 98 ± 1% with the C_{60}-buffered Cu/SiO_2 catalyst (C_{60}-Cu/SiO_2) at ambient H_2 pressure. No substantial limitations in mass and heat transfer were observed in the reaction system at gauge pressure (the pressure above outside ambient) of 1 bar that merely supplies the necessary driving force for substrate diffusion. The catalytic parameters are listed in table S1, and the H_2 pressure for the reaction is equal to gauge pressure plus ambient pressure (the same expression is used throughout this article). An H_2/DMO ratio greater than a stoichiometry ratio is commonly used in industry for a number of purposes, such as removal of excess heat to facilitate the exothermal DMO hydrogenation (ΔG = −31.07 kJ mol^{-1}). After accelerating the reaction conditions by increasing the weight liquid hourly space velocity (WLHSV) from 0.6 $g_{(DMO)}$ $g_{(Catalyst)}^{-1}$ $hour^{-1}$ (simplified as h^{-1} in the following text) at 1 bar to 8.4 $hours^{-1}$ at 30 bar, C_{60}-Cu/SiO_2 maintained high DMO conversion and EG selectivity, whereas at the high WLHSV of 8.4 $hours^{-1}$, the activity of the pristine Cu/SiO_2 catalyst dropped to only a trace yield of EG, and the incomplete hydrogenation product methyl glycolate (MG) was observed (figs. S2 and S3 and table S2).

Lower hydrogenation pressure increased the EG selectivity over the C_{60}-Cu/SiO_2 catalyst, which was 96.2% at 30 bar but 98.0% at 1 bar, and although 20 by-products formed at 30 bar, only 2 by-products were observed at 1 bar (fig. S4). The EG yield at 1 bar is close to the equilibrium limit (98.7%, fig. S5). The apparent activation barrier E_a of C_{60}-Cu/SiO_2 is much lower than that of Cu-SiO_2 by ~54 kJ mol^{-1} (Fig. 1D). The introduction of C_{60} apparently allows the Cu catalyst to activate substrates more efficiently (table S3) (*14*–*19*), and by contrast to those working in high pressures (>20 bar) of H_2 previously reported for the DMO-to-EG process, even in homogeneous pathways (*26*–*28*).

A scale-up experiment (Fig. 1E and fig. S6) with 12.0 g of C_{60}-Cu/SiO_2 was conducted under typical reaction conditions with H_2/DMO = 100 (v/v) and WLHSV = 0.6 $hours^{-1}$. The external mass diffusion cannot be neglected under ambient pressure as the Mears' criterion is >0.15 (table S1). The H_2 pressure was thus set at 3 bar to ensure sufficient substrate diffusion, and as compensation to the pressure, the H_2/DMO ratio was reduced from 200 (v/v) in the initial microscale experiments to 100 (v/v) in the scale-up experiment. For the first 32 hours in the scale-up DMO-to-EG experiment, the temperature was set at 190°C according to the microscale test. However, overhydrogenated by-products (ethanol and butanediols) were produced. The following temperature was set at 182°C with a fluctuation of ±8°C to sustain a high EG yield of >98% up to 1000 hours (Fig. 1E). As extrapolated along the statistics lines of DMO conversion (up to 100%) and EG selectivity (>98%), no decreased yield was observed even after 1000 hours. The spent catalyst can be reused and shows almost no aggregation for the Cu nanoparticles (NPs) therein (fig. S7).

Transmission electron microscopy (TEM), scanning transmission electron microscopy-electron diffraction (STEM-EDX), and line-scan electron energy loss spectroscopy (EELS) were used to establish the morphologic structures of the as-prepared Cu/SiO_2 and C_{60}-Cu/SiO_2 catalysts, as well as identify distributions of Cu and C_{60} in the catalysts (Fig. 2 and figs. S6 and S8 to S10). The samples contained dispersed Cu NPs with sizes ranging from 2 to

[1]State Key Laboratory of Physical Chemistry of Solid Surfaces, National Engineering Laboratory for Green Chemical Production of Alcohols-Ethers-Esters, Collaborative Innovation Center of Chemistry for Energy Materials (iChEM), College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, China. [2]State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, China. [3]Xiamen Funano New Materials Technology Co., Ltd., Xiamen, China.
*Corresponding author. Email: yzyuan@xmu.edu.cn (Y.Z.Y.); syxie@xmu.edu.cn (S.Y.X.)
†These authors contributed equally to this work.

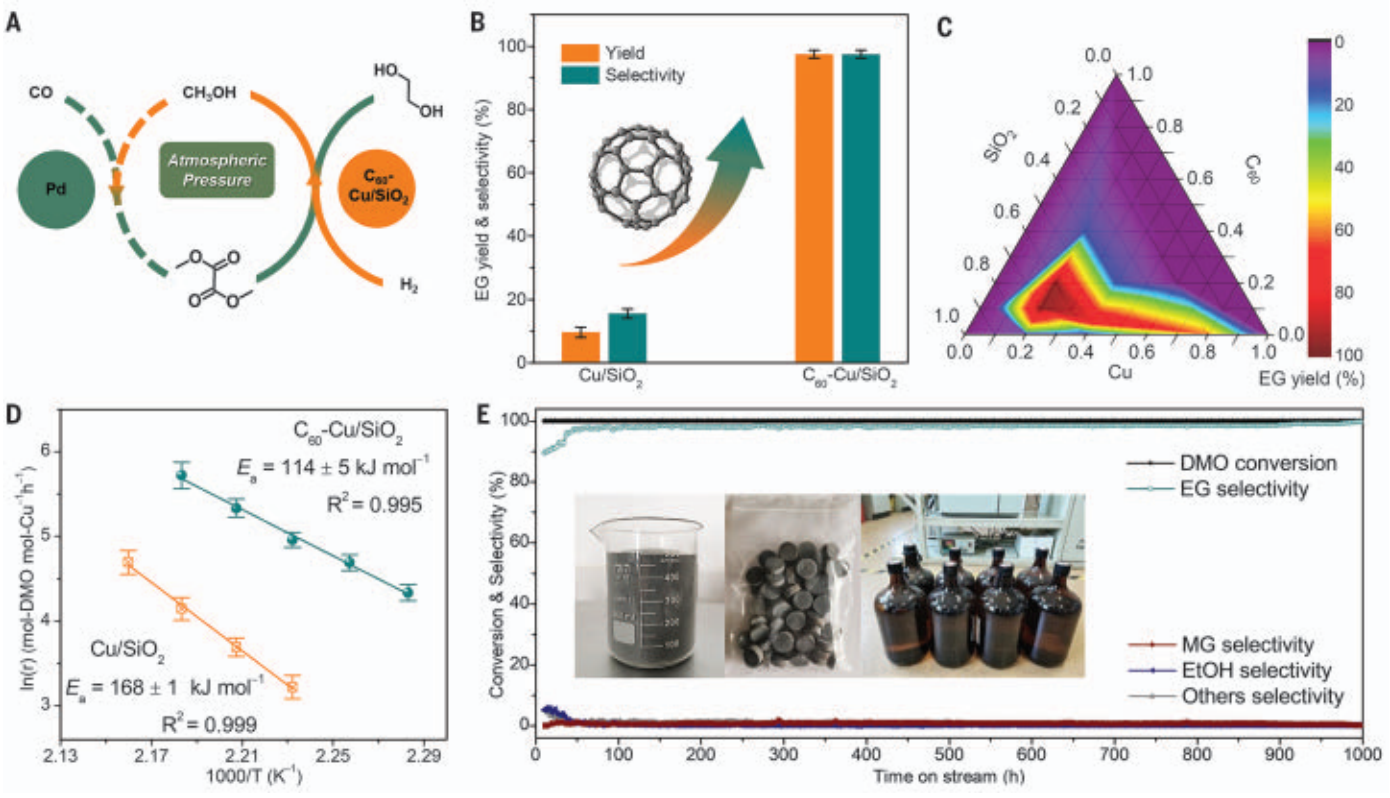


Fig. 1. Catalytical performance of Cu/SiO_2 and C_{60}-Cu/SiO_2 (C_{60}, 10 wt %; Cu, 20 wt %). (**A**) Scheme of the two-step approaches for the synthesis of EG from syngas through DMO hydrogenation. (**B**) Catalytic performance of Cu/SiO_2 improved by C_{60} at 1 bar, H_2/DMO = 200 (v/v, volume ratio), temperature 190°C, WLHSV 0.6 h^{-1}. (**C**) Two-dimensional contour snapshot of catalytic activity over xC_{60}-yCu-$zSiO_2$ ($x + y + z = 1$) with different formulations at 1 bar, H_2/DMO = 200 (v/v), temperature 190°C, WLHSV 0.6 h^{-1}. (**D**) Comparison of activation energy with Cu/SiO_2 and C_{60}-Cu/SiO_2 catalysts. (**E**) Stability tests in DMO hydrogenation with C_{60}-Cu/SiO_2 (scaled up to 12.0 g with the C_{60}-Cu/SiO_2 catalyst) at 3 bar, H_2/DMO = 100 (v/v), WLHSV 0.6 h^{-1}. Insets: catalyst powder (left), pellets (middle), and EG/methanol products (right).

Table 1. Catalytical performance of Cu/SiO_2 and C_{60}-Cu/SiO_2 for hydrogenation. C_{60} and Cu, 10 and 20 wt %, respectively. *P*, pressure; *T*, temperature; Δ*I*, performance improvement by C_{60} on the basis of the increment of yield.

Entry	Substrate	Product	*P*/bar	*T*/°C	WLHSV/h^{-1}	Catalyst	Yield/%	Selec./%	Δ*I*/%
1	Ethyl acetate	Ethanol	1	190	0.60	Cu/SiO_2	NIL	NIL	∞
						C_{60}-Cu/SiO_2	90.1	100.0	
2	Diethyl oxalate	Ethanol	1	240	0.60	Cu/SiO_2	47.7	47.7	80
						C_{60}-Cu/SiO_2	85.9	85.9	
3	Diethyl malonate	1,3-propanediol	30	190	0.60	Cu/SiO_2	8.8	21.4	620
						C_{60}-Cu/SiO_2	63.1	74.2	
4	Dimethyl succinate	butyrolactone	1	230	0.60	Cu/SiO_2	53.4	100.0	71
						C_{60}-Cu/SiO_2	91.6	94.9	
5	Dimethyl maleate	butyrolactone	1	240	0.60	Cu/SiO_2	72.6	100.0	32
						C_{60}-Cu/SiO_2	95.8	95.8	
6	Methyl lactate	1,2-propanediol	1	180	0.18	Cu/SiO_2	39.0	39.0	113
						C_{60}-Cu/SiO_2	83.2	83.2	
7	Methyl pyruvate	Methyl lactate	1	150	0.18	Cu/SiO_2	35.3	67.5	171
						C_{60}-Cu/SiO_2	95.5	95.5	
8	Methyl pyruvate	1,2-propanediol	1	180	0.18	Cu/SiO_2	6.3	8.4	1060
						C_{60}-Cu/SiO_2	72.7	72.7	

5 (nm). N_2O chemisorption in connection with H_2 pulse reduction was adopted to measure the exposed Cu surface area and dispersion (table S2). The Cu surface areas of Cu/SiO_2 and C_{60}-Cu/SiO_2 obtained are similar to each other (31 versus 34 m^2/g), consistent with observations of similar particle sizes (2.9 versus 3 nm) with TEM.

Thermogravimetric analysis in figs. S11 and S12 indicated that 90% of the C_{60} was successfully introduced for a loading of 10 wt % for C_{60}-Cu/SiO_2. Enlarged images of the NPs show C_{60} molecules anchoring on the surface of Cu NPs (Fig. 2, B and D, and fig. S9B). To estimate the distribution of C_{60} across the catalyst, the derivative of the thermogravimetric curve (fig. S12) was analyzed and the peak was deconvoluted at 380° to 540°C, attributed to C_{60} on the Cu surface. The data indicate that most of the C_{60} is loaded on the Cu surface (66% for C_{60}-Cu/SiO_2 with 10 wt % C_{60}) rather than on SiO_2.

X-ray diffraction patterns (XRD) of C_{60}-Cu/SiO_2 show diffraction peaks that can be assigned to C_{60} at 2θ = 10.8°, 17.7°, and 20.8°, corresponding to the (111), (220), and (311) planes, respectively. Additionally, the increasing intensities with progressive addition of C_{60} are shown in Fig. 2F and fig. S8 (*29*). Notably, the addition of C_{60} seems to preserve the Cu NPs surface structure because the physical parameters of metallic loading, NP size, pore structure, and specific surface area did not change

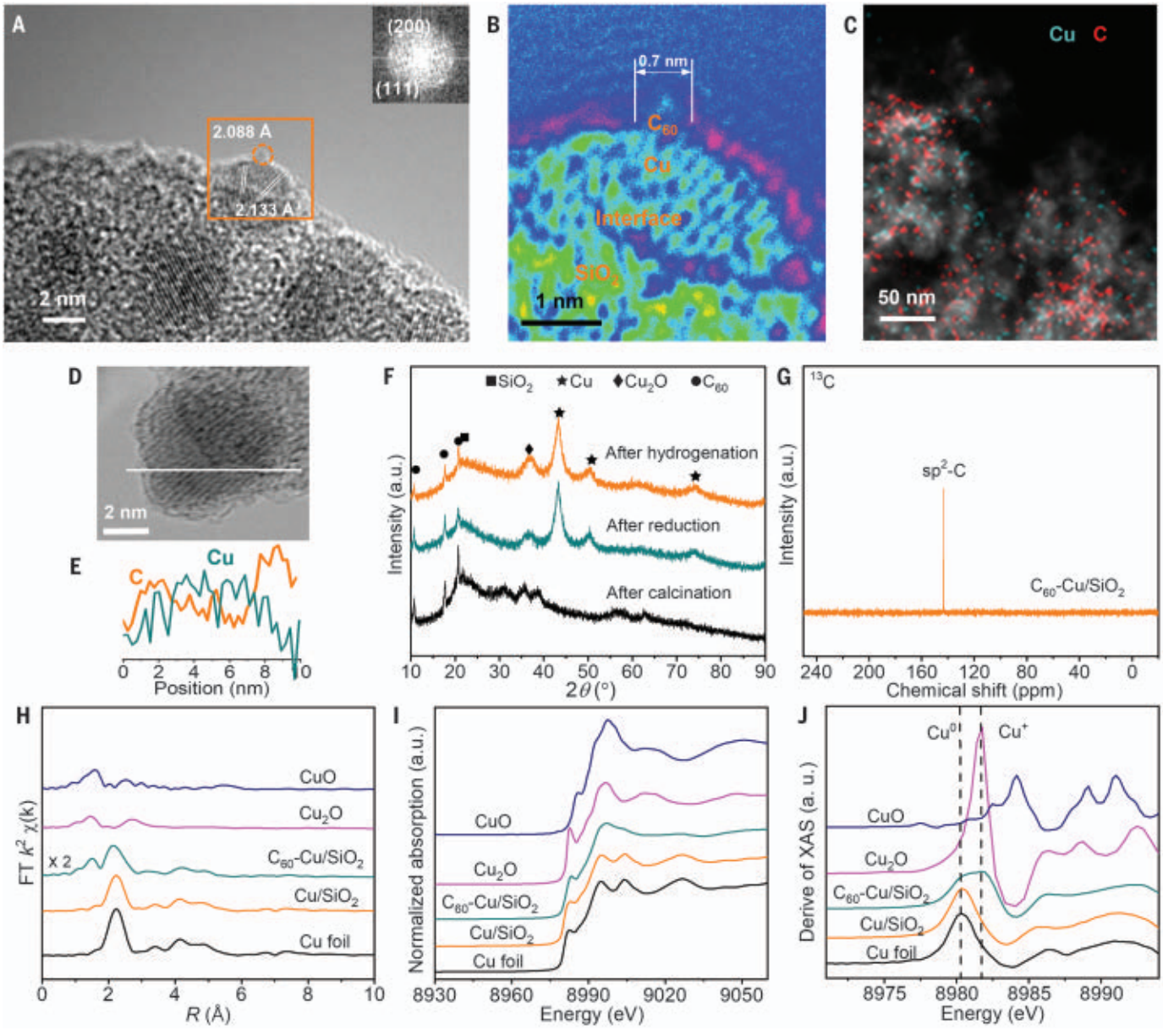


Fig. 2. Geometrical and electronic structural characterizations of C_{60}-Cu/SiO_2 (C_{60}, 10 wt %; Cu, 20 wt %). (**A**) Represented high-resolution TEM image of C_{60}-Cu/SiO_2. Inset: Fast Fourier Transform diffraction pattern of the square area. 2.133 and 2.088 Å are the lattice distances of Cu_2O and Cu species, respectively. (**B**) Color snapshot of the square in (A) showing a plausible C_{60} molecule sitting on a Cu NP. (**C**) Overlap of Cu and C elemental mapping on an SiO_2 matrix from EDX mapping; distribution of Cu is shown by cyan points and C species are shown by red points on the SiO_2 carrier (white frameworks). (**D**) Selected Cu nanoparticle. (**E**) Line EELS profiles in corresponding lines in (D). (**F**) XRD patterns of the C_{60}-Cu/SiO_2 catalyst after calcination, reduction, and hydrogenation reactions. (**G**) ssNMR of ^{13}C for C_{60}-Cu/SiO_2. (**H**) Fourier-transformed magnitudes of the experimental Cu *k*-edge EXAFS spectra. (**I**) Normalized Cu *k*-edge XANES spectra. (**J**) First derivative of Cu *k*-edge XANES spectra over Cu/SiO_2, C_{60}-Cu/SiO_2, CuO, Cu_2O, and Cu foil.

(table S2). Indeed, C_{60} is too big to intercalate into Cu lattice sites but appears to adsorb on the surface of Cu NPs through host-guest interactions. A line scan EELS profile through a Cu NP (Fig. 2, D and E) shows that the typical features of the surface species are carbonaceous and consistent with the spherical shape and subnanometer size of C_{60} adsorbed on Cu NPs. More microscopic images, as well as EELS profiles, are shown in the supplementary materials (figs. S9, B and C, and S10). H_2-temperature–programmed reduction (fig. S13) showed peaks shifting toward higher temperatures with the progressive addition of C_{60}, indicating that C_{60} can inhibit decorated Cu species from reduction.

To probe the chemical environment of the catalysts, solid-state NMR (ssNMR), x-ray absorption fine structure (XAFS), and x-ray absorption near-edge structure (XANES) characterizations were conducted for the catalysts as well as reference samples. As shown in Fig. 2G, only a sharp peak at 143.4 parts per million (ppm) attributed to sp^2-C of fullerene was observed in the ssNMR spectrum without sp^3-C signal (20 to 100 ppm) that would be associated with chemisorption of C_{60} on the Cu NPs. Fourier transform of extended XAFS spectra for C_{60}-Cu/SiO_2 (Fig. 2H and fig. S14) showed a peak at ~1.7 Å that we assigned to electron-deficient Cu species bonding with oxygen (*30*) or C_{60}, in addition to those corresponding to the Cu–Cu bond. Carbonaceous species were dispersed well with Cu in an interplay fashion by forming d–π interactions (fig. S15). The appearance of Cu–C scattering in C_{60}-Cu/SiO_2 implies possible interaction between C_{60} and Cu (table S4). Wavelet transform of XAFS spectra was indicated in fig. S15. Additionally, C_{60}-Cu/SiO_2 has two more lobes at (k, R) = (1.4, 2.2) and (4.0, 1.5), which could be attributed to the scatterings of Cu–C and Cu–O interaction, respectively. The former can be derived from d–π interactions between Cu and C_{60}, in accordance with the NMR without covalent bonding between the Cu and C. The possibility of observed coordinated d–π interactions between a Cu cluster and C_{60} has literature support from a previous crystallographic study (*31*).

XANES spectra in Fig. 2I show that the edge of C_{60}-Cu/SiO_2 falls between those of Cu foil and Cu_2O, demonstrating that Cu is multivalent. Similarly, first-derivative XANES (Fig. 2J) spectra showed a signature of Cu^0 and Cu^+ species for C_{60}-Cu/SiO_2, which was consistent with the result of Fourier transforms of the Cu k-edge XAFS oscillation in Fig. 2H. Cu^0 is the dominant species for the as-reduced Cu/SiO_2, whereas Cu^+ increases in C_{60}-buffered Cu/SiO_2, as evidenced by Auger electron spectroscopy (fig. S16).

To understand the catalytic role of the three constituents (namely C_{60}, Cu, and SiO_2) on

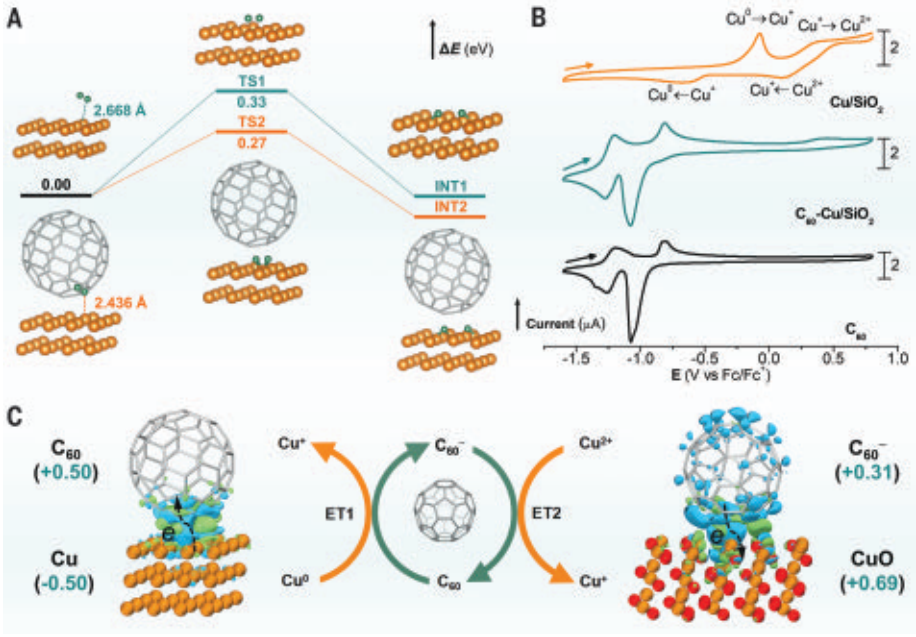


Fig. 3. Electron transfer in a Cu-based catalyst mediated by C_{60}. (**A**) Comparison of H_2 activation with C_{60}-Cu/SiO_2 and Cu/SiO_2 catalysts. TS and INT represent transition state and intermediate, respectively. The green and orange balls represent H and Cu, respectively. The distance between the Cu surface and H_2 is shortened from 2.668 to 2.436 Å when C_{60} is accommodated. (**B**) Cyclic voltammogram of Cu/SiO_2 (upper); C_{60}-Cu/SiO_2 (middle); and C_{60} (bottom) at 0.05 V s^{-1} in propylene carbonate solution containing 0.1 M tetrabutylammonium hexafluorophosphate and 0.016% v/v acetonitrile (segment of four–fifths cycle). All potentials were reported versus the redox couple of the internal ferrocene/ferrocenium (Fc/Fc^+) standard. The potential sweep starts at open circuit potential toward a cathodic direction; C_{60}, 10 wt %; Cu, 20 wt %. (**C**) Calculation results of Bader charges for C_{60}-Cu and C_{60}^--CuO surface interaction systems. The red and orange balls represent O and Cu atoms, respectively. ET represents electron transfer. The plus and minus represent the degree of the Bader charge. The overall Bader charge is 0 and 1 for C_{60}-Cu and C_{60}^--CuO, respectively. The green and blue areas with isosurface contours denote electron accumulation and electron depletion, respectively.

hydrogenation reactions, 20 samples of the C_{60}-Cu/SiO_2 catalysts with variable contents of C_{60}, Cu, and SiO_2 were synthesized and evaluated for the DMO-to-EG process at 1 bar and 190°C with 200 H_2/DMO and 0.6 $hours^{-1}$ WLHSV. As shown in the contour map (Fig. 1C) with EG yield varying as a function of chemical formulation, the highest activity was obtained for fractions between 0.2 to 0.4 for Cu, 0.6 to 0.8 for SiO_2, and 0.05 to 0.25 for C_{60}. Corresponding DMO conversion and EG selectivity have similar high activity compositions (fig. S2, B and C). Discussion of the catalytic roles of C_{60}, Cu, and SiO_2 is detailed in fig. S2 and in the supplementary text. These results imply that Cu is the primary active species for the heterogeneous hydrogenation of DMO and that the catalytic reactivity is sensitive to the addition of C_{60}.

For hydrogenation reactions in Cu-based catalysts, the dissociation of H_2 typically occurs on Cu metallic sites (*20*, *23*). We simulated the activation of H_2 on a crystalline Cu(111) surface and its combination with C_{60} by density functional theory (DFT). As shown in fig. S17, the introduction of C_{60} leads to local electronic redistribution and enhances the local charge density as electrons transfer from Cu to H atom(s) and C_{60} molecules. Meanwhile, the energy barrier for H_2 activation is lower because H_2 obtains more free electrons when coupling with C_{60}–Cu versus Cu alone. Temperature program desorption of H_2 coupled with mass spectrometry (H_2-TPD-MS) has been conducted to investigate the H_2–sorption capacity of the catalysts. As shown in fig. S18, all of the samples show desorption peaks at two regions (60° to 160°C and 300° to 600°C), corresponding to physical and chemical adsorption of H_2, respectively. Clearly, the introduction of C_{60} substantially promotes the chemical adsorption of H_2 as a significantly larger peak at 300° to 600°C. In addition, the bond length of Cu–H is shortened from 2.668 Å to 2.436 Å (Fig. 3A) when C_{60} is accommodated onto the Cu surface, indicating the enhancement of H_2 adsorption. The activation of DMO begins with the nucleophilic attack of adsorbed H atoms to the electron-deficient carbon of the ester group; further, promotion of H_2 activation on the Cu surface by C_{60} can further facilitate the activation of DMO on the Cu surface (*15*).

The theoretical models for C_{60}-Cu and C_{60}^--CuO systems were built to analyze the electron

transfers between C_{60} and Cu^0, C_{60}^- and Cu^{2+}. As shown in Fig. 3C, in the C_{60}-Cu system the Bader charge of Cu (−0.50) is more negative than that of C_{60} (+0.50), implying electron transfer (ET1) from Cu to C_{60}. By contrast, the Bader charge of CuO (+0.69) is higher than that of C_{60}^- (+0.31) in the C_{60}^--CuO system, indicating electron transfer from C_{60}^- to CuO (ET2). Therefore, C_{60} species are beneficial for stabilizing Cu^+.

To mimic electron transfer in dynamic redox processes of DMO hydrogenation over Cu-based catalysts mediated by C_{60}, we used cyclic voltammetry (CV) to transfer electrons to and from the catalyst. Fig. 3B shows the typical CV process of the Cu-based catalysts with or without C_{60} at a scan rate of 0.05 V s^{-1} in an electrolyte containing tetrabutylammonium hexafluorophosphate and propylene carbonate. Two pronounced anodic peaks around −0.2 and +0.3 V were observed for Cu/SiO_2 in the positive scan corresponding to the single-electron oxidation of Cu^0-to-Cu^+ and Cu^+-to-Cu^{2+} respectively. Correspondingly, two peaks in the negative scan around +0.2 and −0.6 V can be assigned to the reversible single-electron reduction of Cu^{2+}-to-Cu^+ and Cu^+-to-Cu^0, as detailed in Fig. 3B and fig. S19, A and B. By contrast, pure C_{60} in the solid state undergoes four redox peaks between −0.5 and −1.6 V (fig. S19C), which are two-step, one-electron–transfer reaction processes as reported previously (*32*).

For C_{60}-Cu/SiO_2, the peaks associated with single-electron oxidation of Cu^0-to-Cu^+ and single-electron reduction of Cu^{2+}-to-Cu^+ were completely absent, which we attributed to modulation of electron transfer by C_{60}. As supported by theoretical calculations for electron transfer between C_{60} and Cu surfaces (ET1, Fig. 3C), the electron lost from Cu^0 does not transfer to the electrode surface but is instead captured by C_{60}, and the current change is undetected in the external circuit. Similarly, the electrons from C_{60}^- can transfer back to Cu^{2+} (ET2, Fig. 3C), which reduced Cu^{2+} to Cu^+ without drawing electrons from the external circuit. Thus, C_{60} species, C_{60} and C_{60}^-, can act as a single-electron acceptor from Cu^0 or donate a single electron to Cu^{2+}, stabilizing Cu^+ and preventing transformation toward Cu^0 or Cu^{2+}.

For C_{60}-Cu/SiO_2, there was one anodic peak around +0.4 V in the positive scan and one cathodic peak around −0.8 V in addition to the peaks of C_{60}. On the basis of a series of electrochemical studies (fig. S19, D and E), we concluded that these two peaks could be assigned to oxidation of Cu^+-to-Cu^{2+} and reduction of Cu^+-to-Cu^0. Both peaks had weaker intensities and emerged at the voltages with higher positive or negative shifts than those of the pristine Cu/SiO_2, implying that it is more difficult to oxidize or reduce Cu^+ species with adsorbed C_{60}. Such a conclusion, regarding C_{60} acting as an electron buffer and creating a more stable environment for electron-deficient Cu species, is further supported by CV tests and DFT calculations on a molecular model of $Cu_{24}O_{24}Si_8R_8$ [R = (2,6-(i-C_3H_7)$_2$$C_6H_3$)N($SiMe_3$)] with a basic unit of Cu^+–O–Si (fig. S20) (*33*).

The improvement of C_{60} as an electron buffer to Cu/SiO_2 catalyst was further investigated for other hydrogenation reactions. As shown in Table 1, Cu/SiO_2 catalysts always show inferior activity compared with C_{60}-Cu/SiO_2. For example, no activity was observed over the Cu/SiO_2 catalyst during hydrogenation of ethyl acetate to ethanol, but the C_{60}-Cu/SiO_2 catalyst exhibited an ethanol yield of 90.1%. For 1,2-propanediol synthesis from methyl pyruvate, the improvement was more than an order of magnitude (6.3% versus 72.7%).

We note that most of the substrates in Table 1 can be derived from biomass, and selective hydrogenation is one of the most viable ways to use biomass. With C_{60}-Cu/SiO_2, the performance can be substantially improved, even under ambient pressure. In addition, C_{60} was recovered quantitatively from the catalysts (table S5). The recovered C_{60} was further confirmed using mass spectrometry (fig S21), indicating C_{60} was stable throughout the thermal process from calcination, reduction, and hydrogenation reactions. We have further explored the C_{60}-Cu/SiO_2 catalyst for electrochemical reduction of CO_2. The introduction of C_{60} to Cu/SiO_2 enhanced the faradaic efficiency of CO, and stability as shown in fig. S22. The excellent electrocatalytic reduction from CO_2 to CO endows the present work meaningful more to extend the CO_2-to-EG process overall at atmospheric pressure. Thus, the ambient-pressure hydrogenation of DMO catalyzed by C_{60}-Cu/SiO_2 reported could be applied to other thermo- and electrocatalytic reactions.

REFERENCES AND NOTES

1. K. Ravindranath, R. A. Mashelkar, *Chem. Eng. Sci.* **41**, 2197–2214 (1986).
2. S. Rebsdat, D. Mayer, "Ethylene glycol" in *Ullmann's Encyclopedia of Industrial Chemistry* (Wiley-VCH, 2000), pp. 531–544.
3. H. Yue, Y. Zhao, X. Ma, J. Gong, *Chem. Soc. Rev.* **41**, 4218–4244 (2012).
4. R.-P. Ye *et al.*, *ACS Catal.* **10**, 4465–4490 (2020).
5. J. F. Knifton, *J. Am. Chem. Soc.* **103**, 3959–3961 (1981).
6. T. Masuda, K. Murata, A. Matsuda, *Bull. Chem. Soc. Jpn.* **59**, 1287–1289 (1986).
7. J. Zheng *et al.*, *J. Phys. Chem. C* **119**, 13758–13766 (2015).
8. H. Miyazaki *et al.*, Ube, Japanese patent 57-180432 (1982).
9. H. Yue, X. Ma, J. Gong, *Acc. Chem. Res.* **47**, 1483–1492 (2014).
10. K. Dong *et al.*, *Nat. Commun.* **7**, 12075 (2016).
11. Z.-N. Xu *et al.*, *ACS Catal.* **3**, 118–122 (2013).
12. S.-Y. Peng *et al.*, *ACS Catal.* **5**, 4410–4417 (2015).
13. C. Wang *et al.*, *J. Catal.* **337**, 145–156 (2016).
14. G. Cui *et al.*, *Appl. Catal. B* **248**, 394–404 (2019).
15. C. Xu *et al.*, *Nat. Commun.* **9**, 3367 (2018).
16. Z. He, H. Lin, P. He, Y. Yuan, *J. Catal.* **277**, 54–63 (2011).
17. Y. Zhao *et al.*, *Ind. Eng. Chem. Res.* **59**, 12381–12388 (2020).
18. L.-F. Chen *et al.*, *J. Catal.* **257**, 172–180 (2008).
19. Y.-N. Wang *et al.*, *Catal. Sci. Technol.* **2**, 1637–1639 (2012).
20. J. Gong *et al.*, *J. Am. Chem. Soc.* **134**, 13922–13925 (2012).
21. Y. Huang *et al.*, *J. Catal.* **307**, 74–83 (2013).
22. M. M.-J. Li *et al.*, *Sci. Rep.* **6**, 20527 (2016).
23. H. Yue *et al.*, *Nat. Commun.* **4**, 2339 (2013).
24. G. Yu, J. Gao, J. C. Hummelen, F. Wudl, A. J. Heeger, *Science* **270**, 1789–1791 (1995).
25. L.-L. Deng, S.-Y. Xie, F. Gao, *Adv. Electron. Mater.* **4**, 1700435 (2018).
26. U. Matteoli, G. Menchi, M. Bianchi, F. Piacenti, *J. Mol. Catal.* **64**, 257–267 (1991).
27. H. T. Teunissen, C. J. Elsevier, *Chem. Commun.* **7**, 667–668 (1997).
28. X. Fang, C. Zhang, J. Chen, H. Zhu, Y. Yuan, *RSC Advances* **6**, 45512–45518 (2016).
29. K. Vimalanathan *et al.*, *Angew. Chem. Int. Ed.* **56**, 8398–8401 (2017).
30. J. Y. Kim, J. A. Rodriguez, J. C. Hanson, A. I. Frenkel, P. L. Lee, *J. Am. Chem. Soc.* **125**, 10684–10692 (2003).
31. S.-Z. Zhan *et al.*, *J. Am. Chem. Soc.* **142**, 5943–5947 (2020).
32. C. Jehoulet, Y. S. Obeng, Y. T. Kim, F. Zhou, A. J. Bard, *J. Am. Chem. Soc.* **114**, 4237–4247 (1992).
33. G. Tan, Y. Yang, C. Chu, H. Zhu, H. W. Roesky, *J. Am. Chem. Soc.* **132**, 12231–12233 (2010).
34. G. Kresse, J. Furthmüller, *Comput. Mater. Sci.* **6**, 15–50 (1996).
35. J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
36. W. Tang, E. Sanville, G. Henkelman, *J. Phys. Condens. Matter* **21**, 084204 (2009).
37. M. J. Frisch *et al.*, Gaussian 09, Revision A.02, (Gaussian Inc., 2016); https://gaussian.com/g09citation/
38. P. B. Weisz, C. D. Prater, *Adv. Catal.* **6**, 143–196 (1954).
39. S. T. Oyama, X. Zhang, J. Lu, Y. Gu, T. Fujitani, *J. Catal.* **257**, 1–4 (2008).
40. Q. Xie, F. Arias, L. Echegoyen, *J. Am. Chem. Soc.* **115**, 9818–9819 (1993).
41. Y. Yang *et al.*, *J. Am. Chem. Soc.* **117**, 7801–7804 (1995).
42. N. Ji *et al.*, *Angew. Chem. Int. Ed.* **47**, 8510–8513 (2008).
43. J. Sun *et al.*, *Sci. Adv.* **4**, eaau3275 (2018).
44. L. R. Zehner, R. W. Lenton, Atlantic Richfield Co., U.S. Patent 4112245 (1978).
45. R.-P. Ye *et al.*, *ACS Catal.* **8**, 3382–3394 (2018).
46. R.-P. Ye *et al.*, *J. Catal.* **350**, 122–132 (2017).
47. A. Satapathy, S. T. Gadge, B. M. Bhanage, *ACS Omega* **3**, 11097–11103 (2018).
48. A. Satapathy, S. T. Gadge, B. M. Bhanage, *ChemSusChem* **10**, 1356–1359 (2017).

ACKNOWLEDGMENTS

Funding: The work was supported by the National Natural Science Foundation of China (21721001, 92061000, 92061204, 21972113, 22171235, 21827801, 21972120, and 21703100), the National Key Research and Development Program of China (2017YFA0206801, 2017YFA0206802, and 2017YFB0307301), and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA21020800). We also thank XAS station (BL14W1) of the Shanghai Synchrotron Radiation Facility. **Author contributions:** J.W.Z. prepared, characterized, and tested the catalysts, and also wrote the manuscript draft. L.L.H., T.Z.Y., and X.F.L. amplified the synthesis. L.L.H. also organized the catalytic tests. Z.C.C and S.W.Y. performed the DFT calculations. C.H.C and K.S. performed the electrochemical study. X.P.D. and X.Y.C. participated in the early experiments. Y.G.Y. designed the scale-up experiments. H.P.Z. synthesized $Cu_{24}O_{24}Si_8R_8$. P.D. conducted CO_2 electrochemical reduction. G.C.G. developed DMO synthesis with Pd catalyst at ambient pressure. C.F.Z. synthesized C_{60} by arc-discharge of graphite. L.S.Z., S.Y.X., and Y.Z.Y. conceived the overall project. All coauthors discussed the data. **Competing interests:** The authors declare no competing interests. A patent has been filed by Xiamen University and Xiamen Funano New Materials Technology Co., Ltd on the findings reported here. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm9257
Materials and Methods
Supplementary Text
Figs. S1 to S22
Tables S1 to S5
References (*34*–*48*)

21 October 2021; accepted 25 January 2022
10.1126/science.abm9257

AEROSOL OPTICS

Amplification of light within aerosol particles accelerates in-particle photochemistry

Pablo Corral Arroyo[1], Grégory David[1], Peter A. Alpert[2], Evelyne A. Parmentier[1], Markus Ammann[2], Ruth Signorell[1]*

Optical confinement (OC) structures the optical field and amplifies light intensity inside atmospheric aerosol particles, with major consequences for sunlight-driven aerosol chemistry. Although theorized, the OC-induced spatial structuring has so far defied experimental observation. Here, x-ray spectromicroscopic imaging complemented by modeling provides direct evidence for OC-induced patterning inside photoactive particles. Single iron(III)–citrate particles were probed using the iron oxidation state as a photochemical marker. Based on these results, we predict an overall acceleration of photochemical reactions by a factor of two to three for most classes of atmospheric aerosol particles. Rotation of free aerosol particles and intraparticle molecular transport generally accelerate the photochemistry. Given the prevalence of OC effects, their influence on aerosol particle photochemistry should be considered by atmospheric models.

Atmospheric aerosol particles are suspensions of solid and liquid particles in air that influence both climate and air quality (*1*). Aerosol and cloud chemistry play a crucial role in the processing of atmospheric particulate matter and are key parts of global atmospheric models (*2*–*6*). Chemical reactions triggered by sunlight in the gas and particle phase have been recognized as a major contributor to the degradation and oxidation of matter in atmospheric aerosols (*7*). Energy- or charge-transfer reactions driven by triplet states (*8*–*10*), photolysis of nitrate and nitrite (*11*), and photolysis of iron carboxylate complexes (*12*, *13*) are examples of atmospherically relevant photochemical processes that involve the particle phase. Photochemical reactions can also be promoted at the surface of atmospheric aerosol particles (*14*–*16*). It has been reported that interface effects and surface charging can cause acceleration of chemical reactions in microdroplets and nanodroplets (*14*, *17*–*20*), observations that have especially sparked interest in the use of microdroplets as new powerful reactors for chemical synthesis (*15*, *16*).

Here, we report another intriguing phenomenon in particle reactions: the influence of optical confinement (OC) on photochemical reactions in aerosol particles (*21*). OC leads to spatial structuring of the light intensity inside the particle [nanofocusing or shadowing; see supplementary materials (SM) section S1] (*21*–*24*) and hence to spatial variations of photochemical rates. For a photochemical reaction step, the overall reaction rate, j, in a particle is given by

$$j = \frac{I(\lambda)\lambda}{hc}\varphi(\lambda)\sigma(\lambda)\int\varepsilon\left(n,k,\vec{r},\lambda\right)C\left(\vec{r}\right)\mathrm{d}\vec{r} \quad (1)$$

where $\vec{r}$ is the position in the particle; the local light-enhancement factor, $\varepsilon\left(n,k,\vec{r},\lambda\right)$, is the ratio of the local light intensity, $I_{\mathrm{p}}\left(\vec{r}\right)$, to the incident light intensity, I; λ is the wavelength of the light; h is Planck's constant; c is the speed of light; n and k are the real and imaginary parts of the complex index of refraction, respectively; φ is the quantum yield; σ is the molecular absorption cross section; and $C\left(\vec{r}\right)$ is the molecular density of the reactant. Shadowing results from strong light absorption (high k values), reducing the average light intensity in the particle. Because shadowing is also present in extended condensed systems (referred to as "bulk"), particle and bulk reaction rates are comparable in this case (*21*). Nanofocusing is a resonance phenomenon that is tied to the spatial confinement by particles. It increases the average light intensity in the particle compared with the incident intensity—that is, on average, $\varepsilon > 1$—thereby accelerating the reaction (Eq. 1) in particles compared with that in bulk systems, where nanofocusing does not occur. The influence of ε on photokinetics and radiation balance can be substantial, but atmospheric models do not usually account for it. The acceleration of photochemical reactions in typical atmospheric aerosol particles is still awaiting a comprehensive evaluation. Although basic research has demonstrated the overall acceleration of photochemical reactions in single aerosol droplets (*21*), the predicted spatial variation of photochemical rates induced by OC effects has not been directly observed or quantitatively constrained within submicron aerosol particles until now.

Here, we directly imaged the local compositional gradients resulting from OC inside single submicrometer-sized particles (Fig. 1). Highly viscous, dried Fe(III)–citrate (FeCit) particles were exposed to ultraviolet (UV) light ($h\nu$, where ν is the photon frequency), resulting in the photoreduction to Fe(II):

$$\mathrm{Fe^{III}Cit} \xrightarrow{h\nu} \mathrm{Fe^{II}(Cit^{\bullet})} \quad (2)$$

Scanning transmission x-ray microscopy coupled with near-edge x-ray absorption fine structure (STXM-NEXAFS) spectroscopy was used to image the temporal evolution of the Fe (III) fraction, α (eq. S2). A quantitative model proves that OC effects are responsible for the

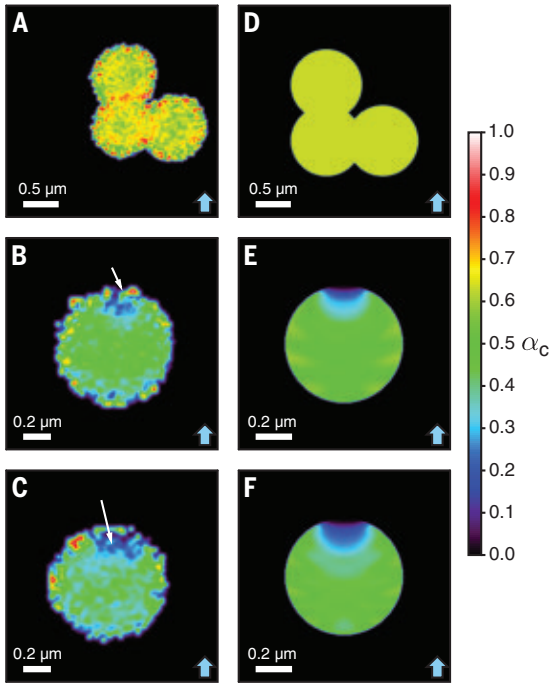


Fig. 1. Maps of the Fe(III) fraction upon UV photoreduction of iron in submicron FeCit particles. (**A** to **F**) STXM-NEXAFS experiments [(A) to (C)] and simulations [(D) to (F)] of the column-averaged Fe(III) fraction, α_c (see color scale bar). α_c is shown before irradiation [(A) and (D)] and after 94 min [(B) and (E)] and 139 min [(C) and (F)] of irradiation with UV light in the direction indicated by the blue arrow. α_c is averaged along the x-ray–beam propagation direction, which is perpendicular to the UV-light propagation direction (fig. S2). The white arrows indicate the hotspot region that is due to nanofocusing.

[1]Department of Chemistry and Applied Biosciences, ETH Zürich, Vladimir-Prelog-Weg 2, CH-8093 Zürich, Switzerland. [2]Laboratory of Environmental Chemistry, Paul Scherrer Institute, 5232 Villigen PSI, Switzerland.
*Corresponding author. Email: rsignorell@ethz.ch

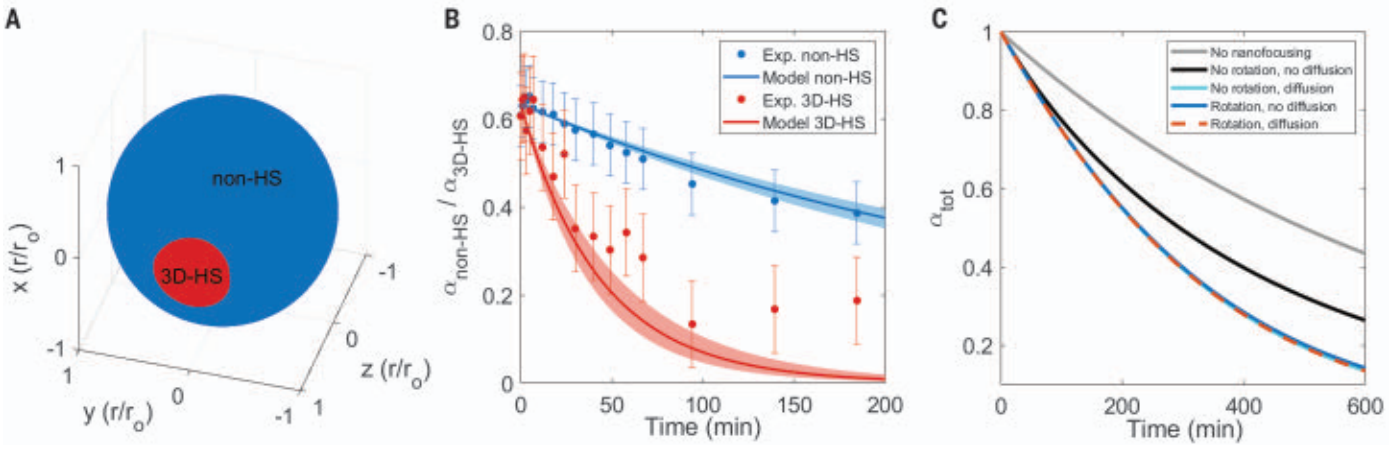


Fig. 2. Influence of nanofocusing, diffusion, and particle rotation on the UV photoreduction of Fe(III) in submicron FeCit particles. (**A**) 3D representation of a particle showing the 3D-hotspot region (red) and the non-hotspot region (blue) (SM section S5). (**B**) Fe(III) fraction in the hotspot region ($\alpha_{3D\text{-}HS}$) and the non-hotspot ($\alpha_{non\text{-}HS}$) region. Circles and curves show experimental data from STXM measurements (SM section S5) and simulations (SM section S7), respectively. The initial Fe(III) ratio was $\alpha(t = 0) = 0.634$. The experimental error bars are either ±0.07 (*28*) or the standard deviations of propagated photon counts, whichever is larger. The shaded regions represent the uncertainty of the model prediction that arises from the experimental uncertainty of the decay rate coefficient in the non-hotspot region ($\beta_{non\text{-}HS}$; SM section S5) and from the uncertainty in the real part of the refractive index (1.5 ± 0.05). (**C**) Photochemical decay of the Fe(III) fraction averaged over the whole particle, $\alpha_{tot}(t)$, calculated for initial iron fractions $\alpha(t = 0) = 1$ and particle radii of 320 nm for five different cases: a particle without rotation and diffusion as in (B) (black trace; SM section S7); a particle without rotation, diffusion, or nanofocusing (gray trace; SM section S8); a particle without rotation but with diffusion (light blue trace; SM section S11); a particle with rotation but no diffusion (dark blue trace; SM section S9); and a particle with rotation and diffusion (dashed orange trace; SM sections S9 and S11).

overall acceleration of the reaction and the spatial inhomogeneity of α that was observed experimentally. Based on this result, we predict how OC effects enhance photochemical reactions in a range of typical atmospheric aerosol particles that are exposed to solar radiation, demonstrating the importance of the phenomenon for the fate of aerosols in the atmosphere.

Figure 1, A to C, shows representative STXM-NEXAFS images of the column-averaged Fe(III) fraction, α_c (see fig. S2A), of FeCit particles acquired before UV irradiation and after 94 and 139 min of UV irradiation ($\lambda \sim 367$ nm), respectively (SM sections S2 to S5). The inhomogeneous spatial depletion of Fe(III) inside the particle that arises from OC effects is clearly visible opposite the side of incidence of the UV light (blue arrow). The white arrows indicate the "hotspot" (Fig. 2A), where the photoreduction is faster than elsewhere in the particle (Fig. 2B). This pattern is the result of nanofocusing, which increases the local light intensity in the hotspot, as confirmed by images simulated for particles with a radius, r_0, of 320 nm (Fig. 1, D to F). The simulations were based on a three-dimensional (3D) particle model that combines light-intensity calculations using the discrete dipole approximation [(*25*) and SM section S6] with a photochemical model for the decay of Fe(III) (SM sections S4 and S7). The results in Fig. 1 are direct observations of the spatial patterning of photochemical reaction products by nanofocusing.

Spatial inhomogeneity was quantified by evaluating the decay of α in two regions within the particle (Fig. 2A and SM section S5): in the 3D hotspot ($\alpha_{3D\text{-}HS}$), where the light intensity was strongly amplified, and in the rest of the particle, which is referred to as the non-hotspot region ($\alpha_{non\text{-}HS}$). The substantially faster decay of $\alpha_{3D\text{-}HS}$ compared with that of $\alpha_{non\text{-}HS}$ (Fig. 2B) highlights the pronounced spatial inhomogeneity of the photoreduction rates caused by OC effects within submicron particles, which is in good agreement with the simulated photochemical decay curves in Fig. 2B (SM section S7). Nanofocusing not only makes the photolysis spatially inhomogeneous but also results in an overall acceleration compared with the reaction in bulk. This becomes evident in the decay of the Fe(III) averaged over the whole particle, α_{tot}, as illustrated in Fig. 2C for simulated decay curves of initially pure FeCit [$\alpha(t = 0) = 1$, where t is the reaction time]. The black decay curve represents the case with nanofocusing, whereas the gray curve corresponds to a hypothetical case without nanofocusing that represents the situation in bulk (SM section S8). In this specific example, the nanofocusing is only moderate but still results in a clear acceleration of the reaction in the particle compared with the reaction in bulk. In other situations, typical of certain atmospheric aerosols, the acceleration is usually even more pronounced (vide infra).

Diffusion time scales of organic molecules within atmospheric secondary organic aerosol (SOA) particles, τ_{mix} (*26*), are compared with a typical photochemical time scale of τ_{photo} = 1 hour in fig. S7. A wide range of diffusivities are included, with an increasing abundance of viscous particles (*27*) (hence with low diffusivity) at higher altitudes. The FeCit particles discussed so far were highly viscous such that diffusion was negligible. For less-viscous atmospheric aerosol particles, however, diffusion needs to be accounted for. To illustrate its effect, Fig. 2C shows the photodecay of a hypothetical FeCit particle with instantaneous diffusion (light blue trace). The comparison with the highly viscous case (black trace) shows that diffusion can further accelerate photochemical reactions when OC effects are present. Diffusion constantly replenishes regions of enhanced light intensity (hotspot) with fresh reactant.

A similar acceleration can result from the free rotation of atmospheric particles, when τ_{photo} is longer than the time scale of rotation. Typical rotation time scales of atmospheric particles in air are shorter than a few hundred milliseconds, whereas atmospheric aerosol photochemistry commonly occurs on time scales, τ_{photo}, from seconds to hours (*7*, *10*, *12*, *13*). Over time, fast particle rotation leads to a better overlap between regions of high reactant concentration and regions of enhanced light intensity (hotspot). Thus, fast rotation amounts to angularly averaging the OC effects. We describe this in terms of a radial light-enhancement factor, ε_r, and a radial Fe(III)

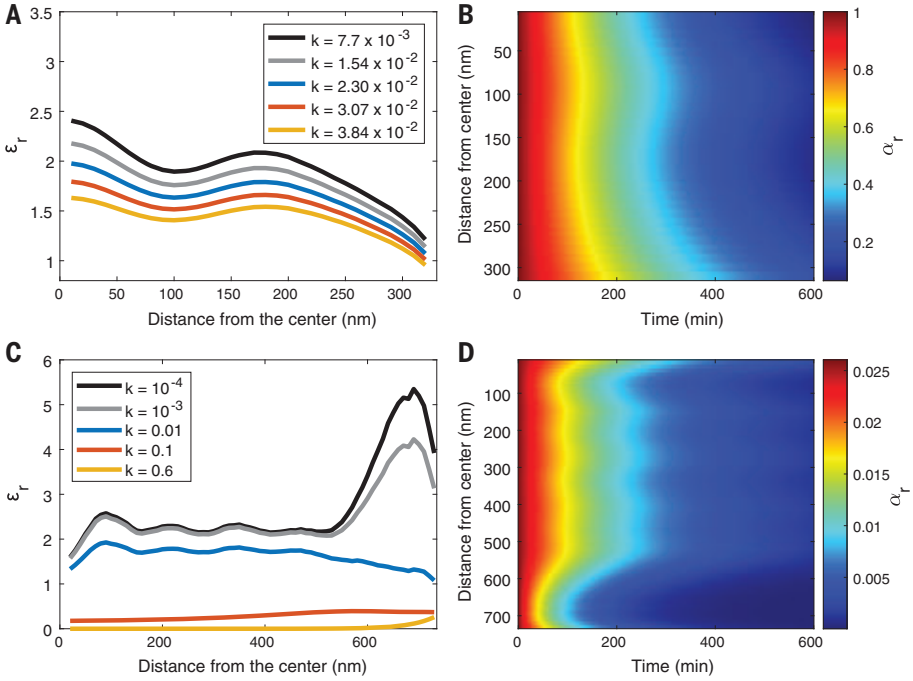


Fig. 3. Radial dependence of light enhancement and Fe(III) fraction in the photoreduction of Fe(III) in rotating FeCit particles. All simulations are for highly viscous (no diffusion), fast-rotating particles for λ = 367.7 nm light, n = 1.5, and φ = 0.016 ± 0.003. (**A** and **B**) Calculations for a fast-rotating particle with radius r_0 = 320 nm and k values that represent experimental conditions (Figs. 1 and 2). Shown in (A) is the radial enhancement factor, ε_r, for selected k values that correspond to experimental Fe(III) fractions of 0.2, 0.4, 0.6, 0.8, and 1, respectively, as a function of the distance from the particle center (0 nm). Shown in (B) are the radial Fe(III) fractions, $\alpha_r(t)$, as a function of the distance from the particle center and the reaction time for $\alpha_r(t = 0) = 1$. ε_r and α_r are obtained by averaging over polar and azimuthal angles (SM section S9). (**C**) Calculations of ε_r for a fast-rotating particle with radius r_0 = 735.4 nm and k values between 10^{-4} (weakly absorbing particle) and 0.6 (strongly absorbing particle). (**D**) Decay of α_r in a rotating particle with radius r_0 = 735.4 nm as a function of time for $\alpha_r(t = 0) = 0.026$.

fraction, $\alpha_r(t)$ (SM section S9). Figure 3A shows ε_r as a function of the distance from the particle center calculated for homogeneous FeCit particles with different Fe(III) concentrations given by the corresponding value of the complex indices of refraction, k. Because the hotspot makes up only a small part of the volume of the particle's outermost layer, angular averaging actually causes ε_r to be smaller near the particle's surface than in its interior. This behavior of ε_r is reflected in the temporal evolution of the radial Fe(III) fraction, $\alpha_r(t)$, in Fig. 3B. α_r decays faster in the interior of the particle than close to its surface, which indicates the persistence of concentration gradients produced by OC effects in viscous particles, in a similar way as observed for the slow reactive uptake of oxygen (*28*). Nevertheless, the decay of Fe(III) averaged over the entire particle, α_{tot} (dark blue trace in Fig. 2C), is still faster than that in the nonrotating particle (black trace). Rotation (dark blue trace) and diffusion (light blue trace) each accelerate the photoreduction to virtually the same extent once they are much faster than the photochemical time scale, as evident from the coincidence of the corresponding traces. In these cases, the acceleration reaches an upper limit that is not surpassed even by simultaneous rotation and diffusion (dashed orange trace). This has an important implication for atmospheric aerosol particles: Because the time scale for particle rotation is typically faster than τ_{photo}, it invariably leads to an acceleration of the photochemistry even if diffusion is slow. Whenever the initial reactant distribution is sufficiently homogeneous, the maximum acceleration is almost reached by rotation alone. Further acceleration by fast diffusion occurs to a limited extent in larger, weakly light-absorbing particles (Fig. 3, C and D).

In cases of a pronounced optical resonance—for example, the excitation of whispering gallery modes (*24*)—the light enhancement near the surface is no longer limited to a single hotspot but instead extends over a large part of the surface layer. Such strong resonances are commonly found in larger particles, as exemplified by a 735-nm particle in Fig. 3, C and D. In contrast to the case of the 320-nm particle (Fig. 3A), the strong enhancement of the light intensity in the surface layer survives the rotational averaging. The result is the pronounced peak in the black and gray traces of ε_r in Fig. 3C. This is again reflected in the temporal evolution of $\alpha_r(t)$ in Fig. 3D: Contrary to the behavior of the smaller particle (Fig. 3B), α_r now decays faster close to the surface of the particle than in the particle's interior. This illustrates the rich variability of the spatial structuring of photochemistry in aerosol particles that is induced by optical resonance effects.

When the light absorption in the particle increases (i.e., increasing k values in Fig. 3C), the light enhancement becomes weaker until the peak near the particle's surface vanishes altogether (blue curve), and ε_r finally drops below one throughout the particle (orange and yellow curves). This is the situation when shadowing dominates (fig. S1B). In this case, there is no acceleration of the photochemistry in particles compared with bulk reactions, which are similarly affected by shadowing (*21*). Such strongly absorbing particles, however, tend to be rare in the atmosphere (see below). We conclude that the resonance effects discussed above generally dominate and accelerate photochemical reactions in particles compared with their bulk counterparts.

With the above results, we assess the influence of OC effects on photochemical reactions in atmospheric aerosol particles. In Fig. 4, we predict the total light-enhancement factor, ε_{tot} (eq. S15), for a range of aerosol particles and conditions relevant to Earth's atmosphere (particle size, refractive index, solar radiation; SM sections S12 and S13). Figure 4A shows the dependence of ε_{tot} on light absorptivity (in terms of k) and particle size (r_0) as the most important parameters. Without loss of generality, we assume spherical particles with a constant n = 1.5. The dashed colored rectangles indicate the typical ranges of absorptivity and particle size for various classes of atmospheric particles, that is, SOA particles (black) (*29*, *30*), humic-like substances (HULIS) particles (blue) (*31*), urban particles (red), rural particles (green) (*32*), soot (brown) (*33*), organic biomass burning particles (purple) (*34*), and sea salt particles (white) (*35*). Calculations for specific substances, accounting for the wavelength dependence of the refractive index, are shown in Fig. 4B for water (*36*); SOA material from limonene, from α-pinene, and from catechol (*29*); and brown carbon (BrC) (*37*). ε_{tot} values less than one (shadowing)—as observed for highly absorbing larger BrC particles—are very rare for atmospheric particles.

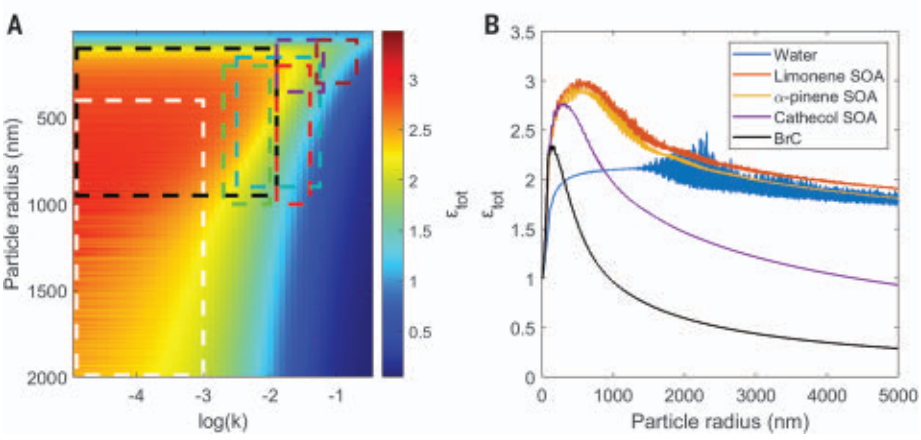


Fig. 4. OC effects in typical atmospheric aerosol particles. (**A**) The total light-enhancement factor, ε_{tot} (see color scale bar), for spherical particles averaged over the wavelength range of 280 to 440 nm as a function of the particle radius (r_0) and the imaginary part of the refractive index (k) for constant n = 1.5 (SM section S12). The dashed colored rectangles indicate the range of properties for different classes of ambient aerosol particles: SOA particles (black), HULIS particles (blue), urban particles (red), rural particles (green), soot (brown), organic biomass burning particles (purple), and sea salt particles (white). (**B**) ε_{tot} as a function of the particle radius for SOA material from limonene, α-pinene, and catechol and for BrC particles. These predictions are for spherical particles using wavelength-dependent refractive index data for the respective materials from the literature in the wavelength range of 280 to 440 nm. For the averaging over the wavelength range of 280 to 440 nm, the relative contribution of each wavelength was weighted according to its relative intensities in sunlight (SM section S12).

Figure 4 shows that enhancement factors for most other atmospheric aerosols lie between two and three, typically reaching their maximum for PM2.5 particles (r_0 < 1.25 μm). The results for water (blue trace in Fig. 4B) show that light-enhancement factors around two are also expected in cloud droplets, certainly up to sizes of many micrometers. Accounting for ε_{tot} in clouds could improve predictions of radical formation in clouds (*38*) and of the formation of aqueous SOAs (*39*).

With typical light-enhancement factors of two to three attained by most atmospheric particles, photochemical reactions in these particles are generally expected to be accelerated on average by the same factor. Accounting for ε_{tot} in SOAs should improve the prediction of the SOA evolution (*40*). This might explain the mismatch between observations and model predictions of SOA mass loss by a factor of two to three, as reported by Hodzic *et al.* (*41*). They used a chemical model to simulate the evolution of submicron α-pinene SOA particles under aging conditions and applied their results to a global chemistry model; however, OC effects were not considered. Given that the authors found photolysis in the particle phase to be a dominant degradation pathway of SOAs, it appears plausible that the neglect of the acceleration of photochemical reactions by light-enhancement effects in the aerosol particles could contribute to the reported discrepancy between model and observation.

Our study provides evidence that light-enhancement effects in typical aerosol particles are more important for photochemical reactions than previously anticipated. For most atmospheric aerosol particles, we argue that an acceleration of photochemical reactions compared with bulk reactions will occur. In view of its prevalence, atmospheric aerosol and cloud models should now account for this phenomenon to improve global chemistry models and climate predictions.

REFERENCES AND NOTES

1. U. Pöschl, *Angew. Chem. Int. Ed.* **44**, 7520–7540 (2005).
2. J. L. Jimenez *et al.*, *Science* **326**, 1525–1529 (2009).
3. M. E. Monge *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6840–6844 (2012).
4. A. Tilgner, P. Bräuer, R. Wolke, H. Herrmann, *J. Atmos. Chem.* **70**, 221–256 (2013).
5. M. Shiraiwa *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11746–11750 (2013).
6. A. Laskin *et al.*, *Science* **301**, 340–344 (2003).
7. C. George, M. Ammann, B. D'Anna, D. J. Donaldson, S. A. Nizkorodov, *Chem. Rev.* **115**, 4218–4258 (2015).
8. S. Canonica, *Chimia* **61**, 641–644 (2007).
9. K. McNeill, S. Canonica, *Environ. Sci. Process. Impacts* **18**, 1381–1399 (2016).
10. P. Corral Arroyo *et al.*, *Environ. Sci. Technol.* **52**, 7680–7688 (2018).
11. H. Herrmann *et al.*, *J. Atmos. Chem.* **36**, 231–284 (2000).
12. C. Weller, A. Tilgner, P. Bräuer, H. Herrmann, *Environ. Sci. Technol.* **48**, 5652–5659 (2014).
13. J. Dou *et al.*, *Atmos. Chem. Phys.* **21**, 315–338 (2021).
14. J. M. Anglada, M. T. C. Martins-Costa, J. S. Francisco, M. F. Ruiz-López, *J. Am. Chem. Soc.* **142**, 16140–16155 (2020).
15. K. J. Kappes *et al.*, *J. Phys. Chem. A* **125**, 1036–1049 (2021).
16. S. Banerjee, E. Gnanamani, X. Yan, R. N. Zare, *Analyst* **142**, 1399–1402 (2017).
17. R. M. Bain, C. J. Pulliam, R. G. Cooks, *Chem. Sci.* **6**, 397–401 (2015).
18. T. Müller, A. Badu-Tawiah, R. G. Cooks, *Angew. Chem. Int. Ed.* **51**, 11832–11835 (2012).
19. P. Nissenson, C. J. H. Knox, B. J. Finlayson-Pitts, L. F. Phillips, D. Dabdub, *Phys. Chem. Chem. Phys.* **8**, 4700–4710 (2006).
20. D. L. Bones, L. F. Phillips, *Phys. Chem. Chem. Phys.* **11**, 5392–5399 (2009).
21. J. W. Cremer, K. M. Thaler, C. Haisch, R. Signorell, *Nat. Commun.* **7**, 10941 (2016).
22. R. Signorell *et al.*, *Chem. Phys. Lett.* **658**, 1–6 (2016).
23. Y. Zhang *et al.*, *NPJ Clim. Atmos. Sci.* **1**, 47 (2018).
24. R. Symes, R. M. Sayer, J. P. Reid, *Phys. Chem. Chem. Phys.* **6**, 474–487 (2004).
25. M. A. Yurkin, A. G. Hoekstra, *J. Quant. Spectrosc. Radiat. Transf.* **112**, 2234–2247 (2011).
26. M. Shiraiwa *et al.*, *Nat. Commun.* **8**, 15002 (2017).
27. A. Virtanen *et al.*, *Nature* **467**, 824–827 (2010).
28. P. A. Alpert *et al.*, *Nat. Commun.* **12**, 1769 (2021).
29. P. Liu, Y. Zhang, S. T. Martin, *Environ. Sci. Technol.* **47**, 13594–13601 (2013).
30. P. F. Liu *et al.*, *Atmos. Chem. Phys.* **15**, 1435–1446 (2015).
31. E. Dinar *et al.*, *Faraday Discuss.* **137**, 279–295 (2008).
32. M. Ebert, S. Weinbruch, P. Hoffmann, H. M. Ortner, *Atmos. Environ.* **38**, 6531–6545 (2004).
33. J. Kim *et al.*, *Aerosol Sci. Technol.* **49**, 340–350 (2015).
34. E. Sarpong, D. Smith, R. Pokhrel, M. N. Fiddler, S. Bililign, *Atmosphere* **11**, 62 (2020).
35. L. Bi *et al.*, *J. Geophys. Res. Atmos.* **123**, 543–558 (2018).
36. G. M. Hale, M. R. Querry, *Appl. Opt.* **12**, 555–563 (1973).
37. J. Li *et al.*, *Atmos. Chem. Phys.* **20**, 4889–4904 (2020).
38. H. Herrmann *et al.*, *Chem. Rev.* **115**, 4259–4334 (2015).
39. B. Ervens, B. J. Turpin, R. J. Weber, *Atmos. Chem. Phys.* **11**, 11069–11102 (2011).
40. M. Shrivastava *et al.*, *Rev. Geophys.* **55**, 509–559 (2017).
41. A. Hodzic *et al.*, *Atmos. Chem. Phys.* **15**, 9253–9269 (2015).
42. P. Corral Arroyo *et al.*, Data Collection: Amplification of light within aerosol particles accelerates in-particle photochemistry. ETH Zürich (2022); https://doi.org/10.3929/ethz-b-000531184.
43. G. David, Registered software: "Particle_internal_intensity_enhancement_factor.py". ETH Zürich (2022); https://doi.org/10.5905/ethz-1007-500.

ACKNOWLEDGMENTS

We acknowledge B. Watts for support during the beamtime at the PolLux end station of the Swiss Light Source and M. Shiraiwa and Y. Li for sharing their global maps of SOA viscosity (fig. S7). **Funding:** This project has received funding from the Swiss National Science Foundation (project 200020_200306) and the European Research Council (Horizon 2020 Research and Innovation Programme grant agreement 786636). The PolLux end station was financed by the German Ministerium für Bildung und Forschung (BMBF) through contracts 05K16WED and 05K19WE2. **Author contributions:** Conceptualization: P.C.A., R.S.; Methodology: P.C.A., G.D., P.A.A., R.S.; Investigation: P.C.A., G.D., P.A.A., E.A.P.; Funding acquisition: R.S.; Project administration: P.C.A., R.S.; Supervision: R.S.; Writing – original draft: P.C.A., G.D., R.S.; Writing – review and editing: P.C.A., G.D., R.S., E.A.P., P.A.A., M.A. **Competing interests:** The authors declare no conflicts of interest. **Data and materials availability:** All data that reproduce the analyses are available in a data repository (*42*). A Python program for the calculations of enhancement factors (SM section S12) is available in a project-associated data archive repository (*43*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm7915
Supplementary Text
Figs. S1 to S8
References (*44–51*)
Movies S1 and S2

12 October 2021; accepted 14 March 2022
10.1126/science.abm7915

CELL BIOLOGY

Epithelial monitoring through ligand-receptor segregation ensures malignant cell elimination

Geert de Vreede†, Stephan U. Gerlach†, David Bilder*

Animals have evolved mechanisms, such as cell competition, to remove dangerous or nonfunctional cells from a tissue. Tumor necrosis factor signaling can eliminate clonal malignancies from *Drosophila* imaginal epithelia, but why this pathway is activated in tumor cells but not normal tissue is unknown. We show that the ligand that drives elimination is present in basolateral circulation but remains latent because it is spatially segregated from its apically localized receptor. Polarity defects associated with malignant transformation cause receptor mislocalization, allowing ligand binding and subsequent apoptotic signaling. This process occurs irrespective of the neighboring cells' genotype and is thus distinct from cell competition. Related phenomena at epithelial wound sites are required for efficient repair. This mechanism of polarized compartmentalization of ligand and receptor can generally monitor epithelial integrity to promote tissue homeostasis.

Epithelial architecture is the fundamental organizing principle of animal tissues. Polarized epithelial sheets provide a contiguous barrier that allows an organ to function in a milieu distinct from the external environment. To maintain the barrier, epithelia must detect threats to their integrity and resolve them. Integrity can be compromised by both physical damage and the production of structurally defective cells. The latter is a frequent feature of oncogenic transformation, and it is important to eliminate such cells before a tumor can form. Deleterious cells can be removed through cell competition, a broadly used mechanism in which "winner" cells of one genotype often induce apoptosis in neighboring "loser" cells (*1*, *2*). In *Drosophila* imaginal discs, cells mutant for the conserved apicobasal polarity regulators *scribble (scrib)* or *discs-large (dlg)* form malignant, "neoplastic" tumors that kill the animal (*3*, *4*). However, prior to tumor growth, small clones of these polarity-deficient cells are efficiently eliminated, allowing a healthy organ to develop. The mechanisms involved have been described as cell competition, during which the *Drosophila* tumor necrosis factor (TNF) ligand Eiger (Egr) binds the TNF receptor (TNFR) Grindelwald (Grnd) in mutant cells, activating the c-Jun N-terminal kinase (JNK) Basket (Bsk), which induces apoptosis [reviewed in (*5*–*7*)]. How polarity loss is coupled to TNF pathway activation to remove oncogenic clones is not known.

We investigated TNF-TNFR interactions during polarity-deficient cell elimination by coculturing imaginal discs ex vivo alongside Egr-Venus (EgrV)–expressing fat bodies (a major endocrine organ) (Fig. 1A) (*8*). EgrV secreted into media associated strongly with clones of *dlg*-depleted cells (Fig. 1, B, C, and E, and fig. S1, A, B, K, and L). Increased EgrV

Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA.
*Corresponding author. Email: bilder@berkeley.edu
†These authors contributed equally to this work.

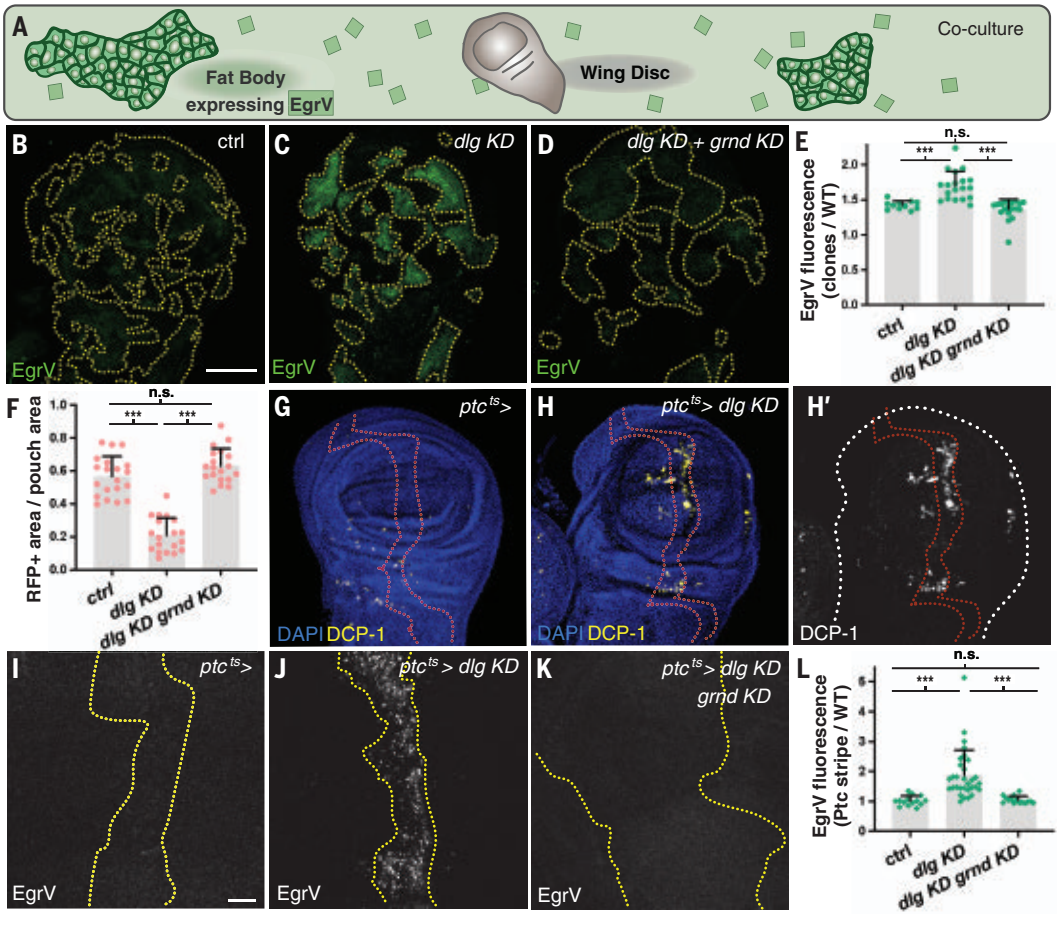


Fig. 1. Heightened binding of Egr to polarity-deficient cells.
(**A**) Diagram of EgrV-expressing fat bodies cocultured ex vivo with imaginal discs. (**B** to **F**) Binding of EgrV and elimination of *dlg*-depleted clones (C) are dependent on cell-autonomous Grnd (D). Control in (B). Dotted lines demarcate clones. Quantitation is shown in (E) and (F). (**G** and **H**) Depletion of *dlg* with *ptc-Gal4* enhances cell death [(H), DCP-1]. Dotted lines indicate stripe of expression. Control in (G). Quantitation is shown in Fig. 2D. (**I** to **L**) EgrV binding, absent in control (I), is increased in the *dlg*-depleted stripe (J), and this requires Grnd (K). Quantitation is shown in (L). Scale bars, 100 μm in (B); 10 μm in (I). Information on statistical tests is provided in the supplementary materials, materials and methods, and table S3.

binding is specific for polarity-deficient elimination; it is seen in *scrib* mutant clones but not loser cells outcompeted by Myc-overexpressing neighbors (fig. S1, G to J). Egr binding, like cell elimination, depends on Grnd (Figs. 1, D to F, and 2I and figs. S1C and S3, C to E). We used *patched-GAL4* (*ptc-GAL4*) to conditionally deplete *dlg*, generating a consistent stripe of apoptotic cells that accumulate EgrV (Fig. 1, G to L, and figs. S1, D and E, and S3, A and B) (*9*, *10*). As in clones, Grnd depletion blocked elimination and led to overgrowth (Fig. 1K and fig. S1F). Mechanical wounding also activates JNK signaling (*11*), and wound sites bind secreted Egr in a Grnd-dependent manner (fig. S2, A to I). Increased Egr-Grnd binding is thus associated with malignant cell elimination and physical wounding, both of which disrupt epithelial integrity.

Because JNK activation in both cases above is associated with Egr binding, we investigated the underlying mechanism. Data argue against increased Grnd levels (fig. S5, A to R), changes in Grnd N-glycosylation (fig. S5S) (*8*), altered endocytic dynamics (fig. S4, A to L) (*12*), or increased Egr levels (fig. S6, A to Q) as mediators of *dlg*-depleted cell apoptosis. In functional experiments, neither autocrine nor paracrine epithelial Egr was required (Fig. 2, A, D to F, H, and L, and fig. S3, F and J). Because polarity-deficient clones survive in an animal completely devoid of Egr (Fig. 2, G and L) (*10*, *12*, *13*), the Egr required for elimination must come from another source.

Both fat body and hemocytes (innate immune cells) produce Egr (*14*–*17*). We codepleted *egr* and *dlg* simultaneously from both the disc stripe and these tissues. Hemocytes did not associate with Dlg-deficient cells, and codepletion of hemocyte Egr had no impact on elimination (Fig. 2, B and D, and figs. S3G and S6, R and S). Codepletion of fat body Egr prevented apoptosis in the stripe; Dlg-deficient cells persisted and overgrew (Fig. 2, C and D, and fig. S3, H to J). *scrib* mutant disc clones persisted upon Egr depletion in fat body and hemocytes (Fig. 2, J and L) but not when Egr was depleted in hemocytes alone (Fig. 2, K and L). Wound healing was also perturbed by depletion of fat body Egr (fig. S2, J and K). Together, these data indicate that circulating Egr is essential for full activation of Grnd and JNK signaling in response to epithelial interruptions.

The above results prompt consideration of ligand and receptor localization in this signaling axis. Fat body–produced Egr is secreted into hemolymph (circulatory fluid), which bathes the disc basolateral surface (Fig. 3A) (*15*). However, steady-state Grnd is apically polarized (Fig. 3, B and C) (*7*). Coculture experiments revealed that EgrV binds only basally, suggesting limited access of circulating Egr to

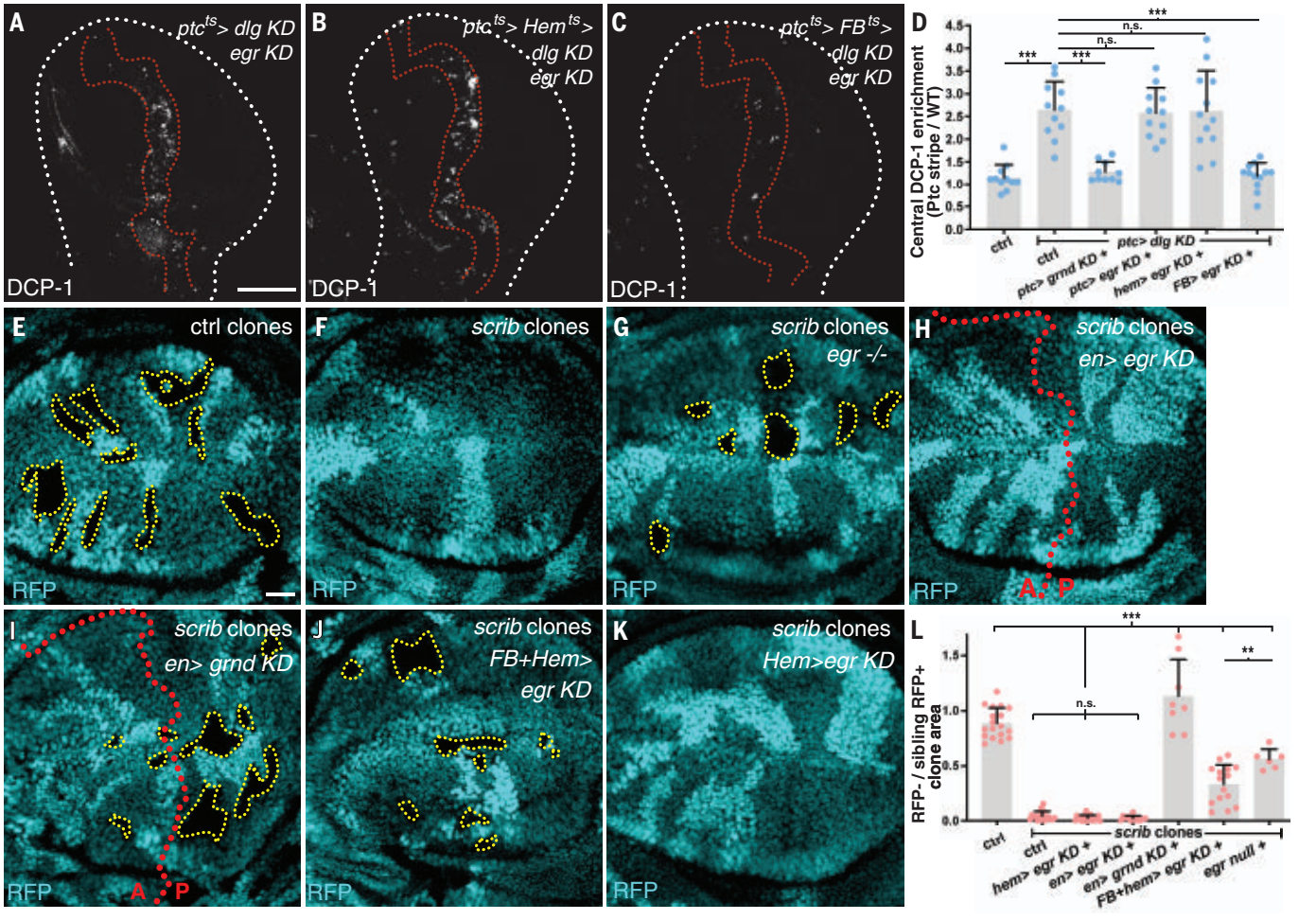


Fig. 2. Egr required for cell elimination derives from circulation. (**A** to **D**) *dlg* cell elimination in the stripe (DCP-1, red dotted lines) is not prevented by codepletion of autonomous *egr* (A) nor *egr* codepletion in hemocytes (B). Depletion of *egr* from fat body and stripe prevents *dlg* cell apoptosis (C). Quantitation of apoptosis shows requirement for autonomous Grnd and fat body–produced Egr (D). (**E** to **L**) Although WT mitotic clones survive (E), *scrib* clones are eliminated (F). *scrib* clones survive in entirely *egr*-mutant animal (G), but are eliminated in a field of *egr*-depleted cells (H). As with autonomous depletion of *grnd* (I), *scrib* clone elimination is blocked when *egr* is depleted in fat body and hemocytes (J) but not in hemocytes alone (K). Quantitation is shown in (L). Dotted lines indicate clone boundaries or posterior (P) compartment gene depletion. Scale bars, 100 μm in (A); 25 μm in (E).

Grnd (Fig. 3C and fig S8, D and E). Dextran assays demonstrated that discs do not display transepithelial permeability, but luminal access can be induced through wounding (Fig. 3, D and E). We tested whether fat body–produced EgrV could bind to an extracellular nanobody (GrabFP) targeted to either polarized epithelial surface in vivo (fig. S8A) (*18*). EgrV bound robustly to basal nanobodies (Fig. 3F) but only slightly to apical nanobodies, and basal signal of these cells was higher (Fig. 3G). After mechanical wounding of the latter discs, EgrV bound apically instead (Fig. 3H). Enhanced EgrV binding was also seen when Dlg-depleted cells mispolarize apical nanobodies (fig. S8, B and C). Thus, TNF ligand and its receptor are normally segregated by the epithelial barrier. However, inducing transepithelial permeability was not sufficient to initiate cell elimination (fig. S7, A to J).

We therefore examined Grnd localization during polarity-deficient cell elimination and found it mispolarized basolaterally (Figs. 3C and 4, A and B, and fig. S8, F and G). This is not due to cell death or basal extrusion (fig. S5, P to R). When these discs are cocultured, bound EgrV is predominantly basal (Fig. 4, A and B). Inhibiting JNK rescued apoptosis but not basal Grnd localization, and again EgrV bound basally (Fig. 4C and fig. S8H). In wounded cells also, EgrV bound predominantly basally, although tissue damage prevented rigorous analysis of Grnd localization (Fig. 4D and fig. S8I). These data suggest that receptor mispolarization allows access to basally circulating ligand, triggering JNK activation and adaptive homeostatic responses, including cell elimination and wound-healing.

Elimination of polarity-deficient cells has long been described as a form of cell competition, albeit regulated by pathways distinct from *Minute* or Myc competition (*1*, *2*, *5*–*7*). A defining feature of cell competition is that elimination of loser cells requires neighboring winners of a different genotype. Yet circulating Egr can access basolaterally mislocalized Grnd in any polarity-deficient cell, regardless of its neighbor. We therefore reconsidered the requirement for wild-type (WT) cells in death of *scrib*-class mutant cells.

We asked whether polarity-deficient discs containing no WT cells showed the same dependence on circulating Egr as that of polarity-deficient cells with WT neighbors. *scrib* discs bound EgrV specifically at their hemolymph-contacting periphery, paralleling the activation of JNK reporters (Fig. 4, E and F, and fig. S9, A and B) (*19*). Apoptosis in *scrib* discs is also enriched peripherally, compared with the "core," which lacks EgrV binding (Fig. 4J). Peripheral apoptosis and JNK activation were normalized when circulating Egr was depleted, and *scrib* discs were larger, which is consistent with a hypothesis that multilayered tissue architecture allows some *scrib* cells to evade hemolymph Egr and overproliferate to form tumors (Fig. 4, F to L) (*19*). *scrib* disc periphery mitotic rates were increased in fat body Egr–depleted animals, likely due to relief of JNK-mediated cell-cycle stalling (fig. S9, C to E) (*20*). Thus, the same mechanisms that eliminate small clones of polarity-deficient cells also kill polarity-deficient cells and limit their growth in a noncompetitive situation. These results challenge the paradigm that elimination of *scrib* cells is due to classical cell competition and suggest that the mechanism we describe is a distinct pathway that couples epithelial organization to tissue homeostasis.

All epithelia need to monitor their integrity and respond when breaches are detected. Because most tumors arise in epithelial tissues,

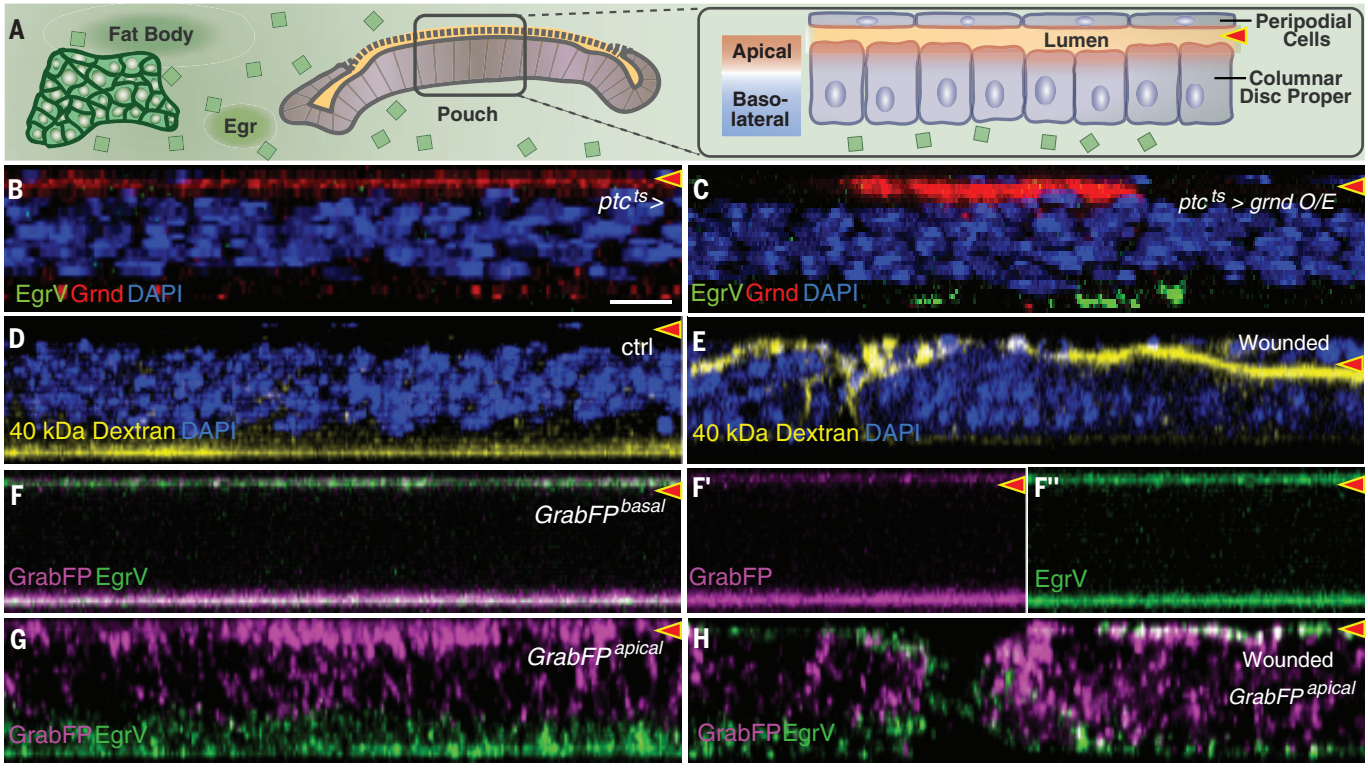


Fig. 3. Egr binds basolaterally to polarity-deficient cells. *X-Z* cross sections show disc proper below and peripodium above. Lumen is indicated with red arrowheads. (**A**) Diagram showing relationship of disc epithelial barrier to hemolymph. Egr secreted by fat body bathes the basolateral surface but is excluded from apical surface and lumen. (**B** and **C**) Grnd is apically localized (B), even when overexpressed (C), but bound EgrV is exclusively basolateral. (**D** and **E**) Dextran in media is excluded from lumen of intact discs (D) but can enter wounded discs (E). (**F** to **H**) Basolateral GrabFP binds strongly at basal surface to EgrV produced by fat bodies (F). Signal at top is peripodial basal surface. [(**F′** and **F″**)] Left half of (F). Apical GrabFP binds EgrV only at the basolateral surface (G), but wounding enables strong apical binding of EgrV as well (H). Scale bar, (B) 10 μm.

preventing the growth of malignant clones within them must also be a priority. In this work, we show a mechanism for tumor elimination that arises from an intrinsic property of the epithelial barrier: its ability to compartmentalize a luminal environment segregated from the external milieu. *Drosophila* TNF circulates systemically and bathes basal organ surfaces but is latent because of TNFR's strict apical localization. However, when neoplastic cells arise, their altered polarity induces basal localization of TNFR, where it binds ligand and triggers apoptotic signaling. A similar axis promotes wound healing; if the epithelium is physically ruptured, ligand can meet receptor and contribute to a pro-healing JNK signaling program. Thus, a common molecular mechanism inherent to epithelial geometry—a mechanism that recognizes polarity changes as a damage-associated molecular pattern (DAMP) (*21*)—underlies both homeostatic programs (Fig. 4M). The function of the TNF-TNFR system described here as an in vivo sensor of epithelial integrity raises the possibility

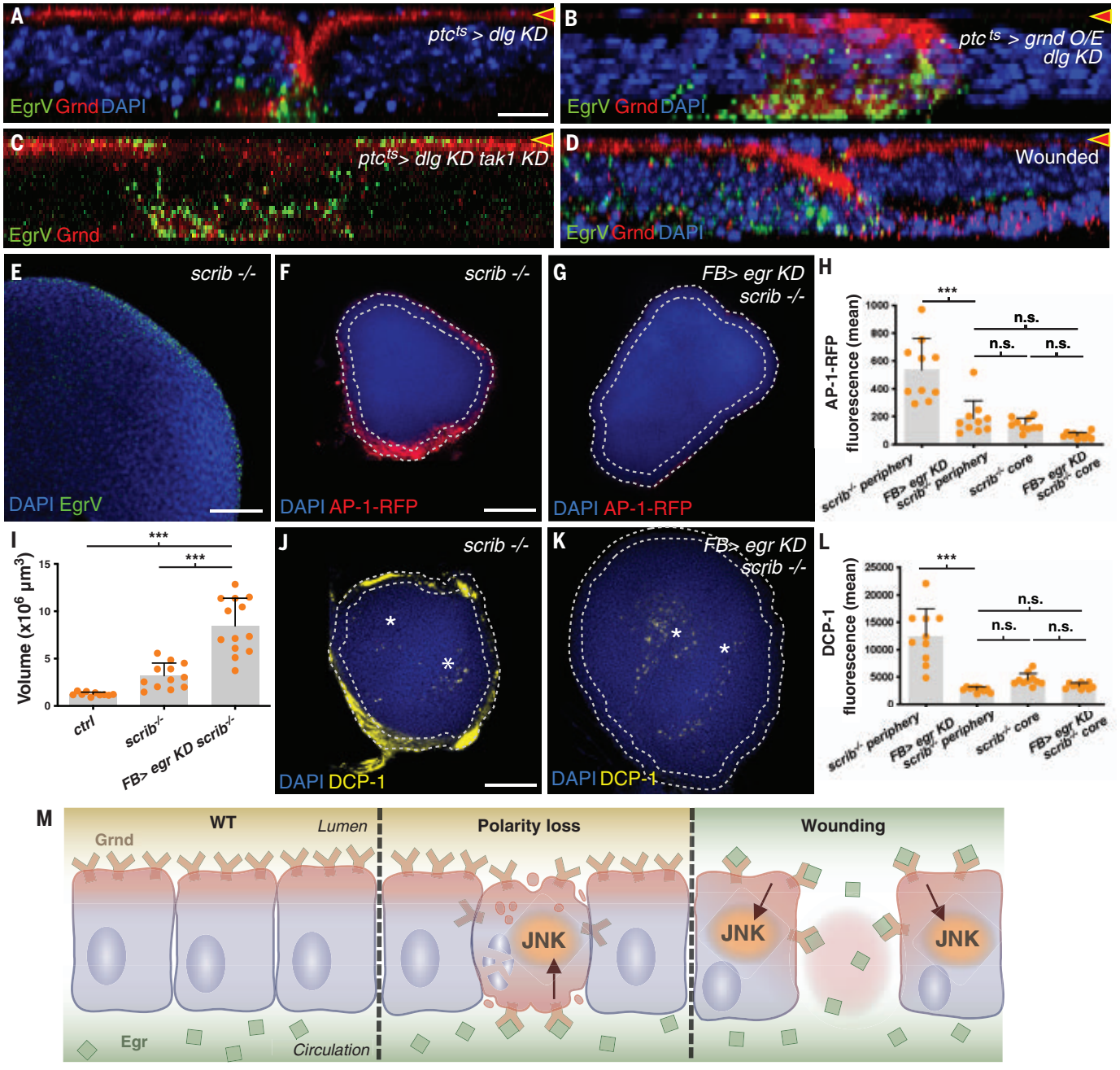


Fig. 4. Mispolarization of Grnd permits Egr binding and cell elimination. (**A** to **D**) *X-Z* sections as in Fig. 3. *dlg*-depleted cells mislocalize Grnd basolaterally (A), evident especially when Grnd is overexpressed (B). EgrV binds basally to *dlg*-depleted cells. Codepletion of *dlg* and *tak1* blocks cell elimination but polarity defects remain (C). Mispolarized Grnd binds EgrV at basal surface. Wounded discs show altered Grnd localization and basal EgrV binding (D). (**E** to **L**) *X-Y* sections. Wing discs containing only *scrib* mutant cells bind Egr preferentially at periphery (E). JNK signaling is increased in periphery compared with core (F) and is dependent on circulating Egr (G). Quantitation is shown in (H). *scrib* discs grow larger when circulating Egr is depleted (I), and peripheral apoptosis [(J), DCP-1] is reduced (K). Asterisks indicate DCP-1^{+} cells in core. Quantitation is shown in (L). (**M**) Model for role of polarized segregation of TNF ligand and receptor in epithelial homeostasis. Scale bar, (A) 10 µm; (E) 50 µm; (F) and (J) 100 µm.

that ligand-receptor segregation (*22*) may be a theme general to epithelial maintenance.

REFERENCES AND NOTES

1. L. A. Johnston, *Cold Spring Harb. Perspect. Med.* **4**, a014274 (2014).
2. E. Madan, R. Gogna, E. Moreno, *Curr. Opin. Cell Biol.* **55**, 150–157 (2018).
3. D. Bilder, *Genes Dev.* **18**, 1909–1925 (2004).
4. R. Stephens *et al.*, *J. Mol. Biol.* **430**, 3585–3612 (2018).
5. G. Morata, M. Calleja, *Semin. Cancer Biol.* **63**, 19–26 (2020).
6. R. Nagata, T. Igaki, *Dev. Growth Differ.* **60**, 522–530 (2018).
7. D. S. Andersen *et al.*, *Nature* **522**, 482–486 (2015).
8. G. de Vreede *et al.*, *Dev. Cell* **45**, 595–605.e4 (2018).
9. C.-C. Yang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 1785–1790 (2015).
10. J. B. Cordero *et al.*, *Dev. Cell* **18**, 999–1011 (2010).
11. A. Repiso, C. Bergantiños, M. Corominas, F. Serras, *Dev. Growth Differ.* **53**, 177–185 (2011).
12. T. Igaki, J. C. Pastor-Pareja, H. Aonuma, M. Miura, T. Xu, *Dev. Cell* **16**, 458–465 (2009).
13. C. L. Chen, M. C. Schroeder, M. Kango-Singh, C. Tao, G. Halder, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 484–489 (2012).
14. F. Parisi, R. K. Stefanatos, K. Strathdee, Y. Yu, M. Vidal, *Cell Rep.* **6**, 855–867 (2014).
15. N. Agrawal *et al.*, *Cell Metab.* **23**, 675–684 (2016).
16. M. Muzzopappa, L. Murcia, M. Milán, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7291–E7300 (2017).
17. C. E. Fogarty *et al.*, *Curr. Biol.* **26**, 575–584 (2016).
18. S. Harmansa, I. Alborelli, D. Bieli, E. Caussinus, M. Affolter, *eLife* **6**, e22549 (2017).
19. T. Ji *et al.*, *Dis. Model. Mech.* **12**, dmm040147 (2019).
20. A. Cosolo *et al.*, *eLife* **8**, e41036 (2019).
21. G. Y. Chen, G. Nuñez, *Nat. Rev. Immunol.* **10**, 826–837 (2010).
22. P. D. Vermeer *et al.*, *Nature* **422**, 322–326 (2003).

ACKNOWLEDGMENTS

We thank S. Yoo and C. Liu for help with wounding experiments. **Funding:** This work was funded by NIH grants GM090150 and GM130388 to D.B. and an Independent Research Fund Denmark fellowship 0131-00010B to S.U.G. **Author contributions:** G.d.V., S.U.G., and D.B. designed the research and wrote the manuscript; G.d.V. and S.U.G. conducted experiments and analyzed data. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Materials are available upon request.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abl4213
Materials and Methods
Figs. S1 to S9
Tables S1 to S3
References (*23*–*34*)
MDAR Reproducibility Checklist

18 July 2021; accepted 15 March 2022
10.1126/science.abl4213

HYDROGELS

Hydrogel-based strong and fast actuators by electroosmotic turgor pressure

Hyeonuk Na[1]†, Yong-Woo Kang[1]†, Chang Seo Park[1]†, Sohyun Jung[2]†, Ho-Young Kim[2]*, Jeong-Yun Sun[1,3]*

Hydrogels are promising as materials for soft actuators because of qualities such as softness, transparency, and responsiveness to stimuli. However, weak and slow actuations remain challenging as a result of low modulus and osmosis-driven slow water diffusion, respectively. We used turgor pressure and electroosmosis to realize a strong and fast hydrogel-based actuator. A turgor actuator fabricated with a gel confined by a selectively permeable membrane can retain a high osmotic pressure that drives gel swelling; thus, our actuator exerts large stress [0.73 megapascals (MPa) in 96 minutes (min)] with a 1.16 cubic centimeters of hydrogel. With the accelerated water transport caused by electroosmosis, the gel swells rapidly, enhancing the actuation speed (0.79 MPa in 9 min). Our strategies enable a soft hydrogel to break a brick and construct underwater structures within a few minutes.

Because of their softness, transparency, biocompatibility, and responsiveness to stimuli, hydrogels are attractive candidates for soft actuators and have been applied in various fields such as soft robotics (*1*, *2*), tunable optics (*3*), fluidics (*4*), and biomedicine (*5*). The actuation mechanism of most hydrogel actuators is swelling, driven by osmotic pressure in response to external stimuli such as solvent (*6*), temperature (*7*), pH (*8*), electric field (*9*), and light (*10*). However, hydrogel soft actuators generally suffer from small actuation stress (i.e., actuation force per unit area) and low speed. The weak actuation stress comes from the low elastic modulus and strength of the hydrogel, whereas the low actuation speed is attributed to osmotic swelling, which proceeds with the diffusion of water.

Although the actuation stress of a hydrogel actuator is usually weak (i.e., 1 to 100 kilopascals), the osmotic pressure of the hydrogel—the origin of the actuation—is able to reach ~50 megapascals (MPa) (*11*). Previous studies on hydrogel actuators have not focused on converting the high osmotic pressure to a corresponding strong actuation stress. However, in nature, plant cells harness their high osmotic pressure to achieve strong turgor pressure. Turgor pressure is the hydrostatic pressure in plant cells resulting from the osmosis-driven swelling confined by the stiff semipermeable cell walls. The turgor pressure equalizes with the high osmotic pressure, allowing soft plant cells to support their bodies, dig deep into the soil, and even break solid rocks. Likewise, a system that uses turgor pressure is expected to improve the actuation stresses of the hydrogel actuators by exploiting the high osmotic pressure.

Most hydrogel actuators suffer from low speeds due to their diffusion-based actuation mechanisms. Many studies have attempted to improve these actuation speeds by increasing diffusion rates or using different actuation mechanisms; these include the molecular engineering of stimuli-responsive hydrogels (*12*–*15*), the incorporation of active materials into hydrogel matrices (*16*), use of the elastic potential energy of the hydrogel network (*17*), and the pneumatic or hydraulic actuation of the hydrogel cover structure (*18*). Despite substantial improvements in speed, actuation forces have been limited to <1 N (*19*) because hydrogels are intrinsically soft. A mechanism that accelerates swelling is expected to create synergy with a system that can use turgor pressure and thus is promising in achieving both fast and strong actuations. Swelling can be accelerated by adopting a fast water-transport mechanism. Electroosmosis induces a constant and rapid water flow through a charged porous material under an electric field and can thus accelerate the swelling of hydrogels because the migrating ions drag the nearby water into the charged polymer network.

We report a hydrogel-based actuator with a design that uses turgor pressure and electroosmosis, which can achieve much higher actuation force in a shorter time than conventional hydrogel actuators. Wrapped in a membrane, a hydrogel converts its inherent high osmotic pressure to a large actuation stress. The actuation stress of a hydrogel actuator (bare) can be characterized by the blocking stress (σ_{block}) of a constrained swelling between two rigid plates at the equilibrium state (Fig. 1, A and B); the blocking stress can be predicted with the ideal elastomeric gel model (*20*). The model presents equations of state for a swollen hydrogel (Eqs. 1 and 2), which are used to determine the swelling stress (σ_{sw})—the stress that a hydrogel applies outward—on the basis of the difference between two competing terms: (i) the water inflow pressure (P_{in}), which is the driving force for the volumetric expansion of a hydrogel and is equal to the osmotic pressure (π_{os}) in an aqueous environment ($P_{in} = \pi_{os}$); and (ii) the elastic stress (σ_{el}),

[1]Department of Material Science and Engineering, Seoul National University, Seoul 08826, Republic of Korea. [2]Department of Mechanical Engineering, Seoul National University, Seoul 08826, Republic of Korea. [3]Research Institute of Advanced Materials (RIAM), Seoul National University, Seoul 08826, Republic of Korea.
***Corresponding author. Email: hyk@snu.ac.kr (H.-Y.K.); jysun@snu.ac.kr (J.-Y.S.)**
†These authors contributed equally to this work.

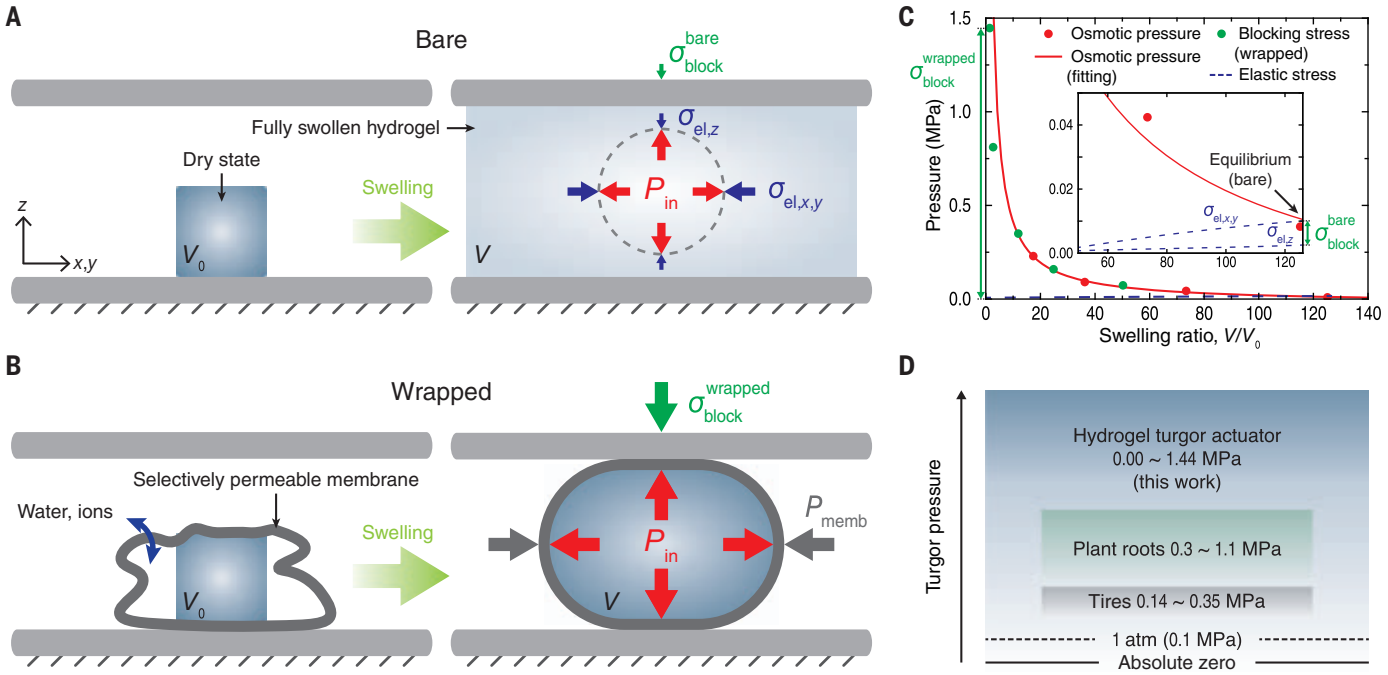


Fig. 1. Design and principle of a hydrogel turgor actuator. (**A** and **B**) Schematic illustrations of constrained osmotic swelling of hydrogel between two fixed plates (A) without membrane (bare) and (B) with membrane (wrapped). The initial hydrogel is a dry polymer cube with a volume of V_0. V, volume of the hydrogel after swelling. Pressure and stress within the gels are represented as colored arrows; red is water inflow pressure (P_{in}), and blue is elastic stress (σ_{el}). The membrane stress (σ_{memb}) in the membrane cross section is represented by dark gray arrows and the blocking stress (σ_{block}) is represented by green arrows. When the hydrogel swells in a membrane by osmosis, the water inflow pressure becomes the osmotic pressure of the gel $P_{in} = \pi_{os}$, and the high osmotic pressure can be sustained by the turgor pressure originating from the membrane tension leading to the generation of a large blocking stress. (**C**) The osmotic pressure and the elastic stress developed in PSPA hydrogel as a function of the swelling ratio (V/V_0). The osmotic pressure of the gel (red dot and line) was measured by dynamic mechanical analysis. The blocking stresses of the hydrogel turgor actuators (green dot) were measured by a universal testing machine and were consistent with the osmotic pressure curve. The inset shows a magnified graph at a high swelling ratio. The blocking stresses of the wrapped gel and the bare gel are marked. The blocking stress of the wrapped gel with a low swelling ratio (1.53) was 1.44 MPa, whereas the osmotic pressure of the bare gel with the same polymer weight was 0.008 MPa. (**D**) Comparison of turgor pressures between our turgor actuator, plants, and tires.

which is the elastic restoring stress of the polymer network which prevents swelling

$$\sigma_{sw,i} = P_{in} - \sigma_{el,i} = \pi_{os} - \sigma_{el,i} \, (i = x, y, z) \quad (1)$$

$$\sigma_{el,i} = \frac{NkT}{J}\left(\lambda_i{}^2 - 1\right)(i = x, y, z) \quad (2)$$

where N is the number of polymer chains per unit volume of a dry network, and kT is the temperature in the unit of energy. π_{os} is a function of the swelling ratio (J) and has the same value regardless of the direction. $J = \lambda_x\lambda_y\lambda_z$, where λ_i is the stretch ratio from the initial hydrogel unit cube in the i direction.

A constrained swelling of a bare gel meets $\sigma_{sw,x} = \sigma_{sw,y} = 0$ and $\lambda_x = \lambda_y$ in the equilibrium state, as shown in Fig. 1A. From Eq. 1, $\pi_{os} = \sigma_{el,x} = \sigma_{el,y}$, and the blocking stress is calculated as

$$\sigma_{block}^{bare} = \sigma_{sw,z} = \pi_{os} - \sigma_{el,z} = \sigma_{el,x} - \sigma_{el,z} = \frac{NkT}{J}\left(\lambda_x{}^2 - \lambda_z{}^2\right) \quad (3)$$

where $NkT/J^{1/3}$ is the shear modulus of a gel, which is generally low (e.g., below 50 kPa) (*20*, *21*), resulting in the weak blocking stress presented by bare hydrogels. The shear modulus of the fully swollen hydrogel used in our experiment was measured as 8.24 kPa with $J = 125$ (Fig. 1C); further, the blocking stress of the bare hydrogel was limited even below NkT, i.e., 41.2 kPa (see methods for details). At the equilibrium state, the remaining osmotic pressure is as low as its weak elastic stress, which leads to the weak blocking stress.

In our turgor actuator (Fig. 1B), which is a hydrogel wrapped in a selectively permeable and relatively stiff membrane, the membrane prevents the swelling of the gel. As a result, a high osmotic pressure can be retained in the hydrogel, thus causing the contribution of the polymer elastic stress (σ_{el}) to the swelling stress (σ_{sw}) to become negligible (Fig. 1C). Then, the swelling stress equals the osmotic pressure:

$$\sigma_{sw,i} = P_{in} - \sigma_{el,i} = \pi_{os} - \sigma_{el,i} \simeq \pi_{os} \, (i = x, y, z) \quad (4)$$

Therefore, the gel inside the membrane can act similar to a liquid under a hydrostatic pressure equal to the osmotic pressure. In the transverse (x, y) direction, the swelling stress is balanced with the membrane pressure (P_{memb}) (see supplementary text and fig. S1).

$$P_{memb} \simeq \pi_{os} \quad (5)$$

In the longitudinal (z) direction, the swelling stress is balanced with the blocking stress of the turgor actuator. From Eq. 4, the blocking stress can be expressed as

$$\sigma_{block}^{wrapped} = \sigma_{sw,z} \simeq \pi_{os} \quad (6)$$

By having the gel constrained in a membrane, the hydrogel turgor actuator can exploit the high osmotic pressure (π_{os}) of the hydrogel as the blocking stress $\left(\sigma_{block}^{wrapped}\right)$.

To verify Eq. 6 experimentally, the blocking stresses of the hydrogel turgor actuators were compared with the osmotic pressures in various swelling ratios (Fig. 1C and fig. S2; see methods for details) (*21*). As shown in Fig. 1C, the blocking stresses are consistent with the osmotic pressure curve (table S1). The actuators achieved a broad range of pressure, up to 1.44 MPa (Fig. 1D).

To evaluate the turgor effect, the actuation stresses and speeds of the bare hydrogel and

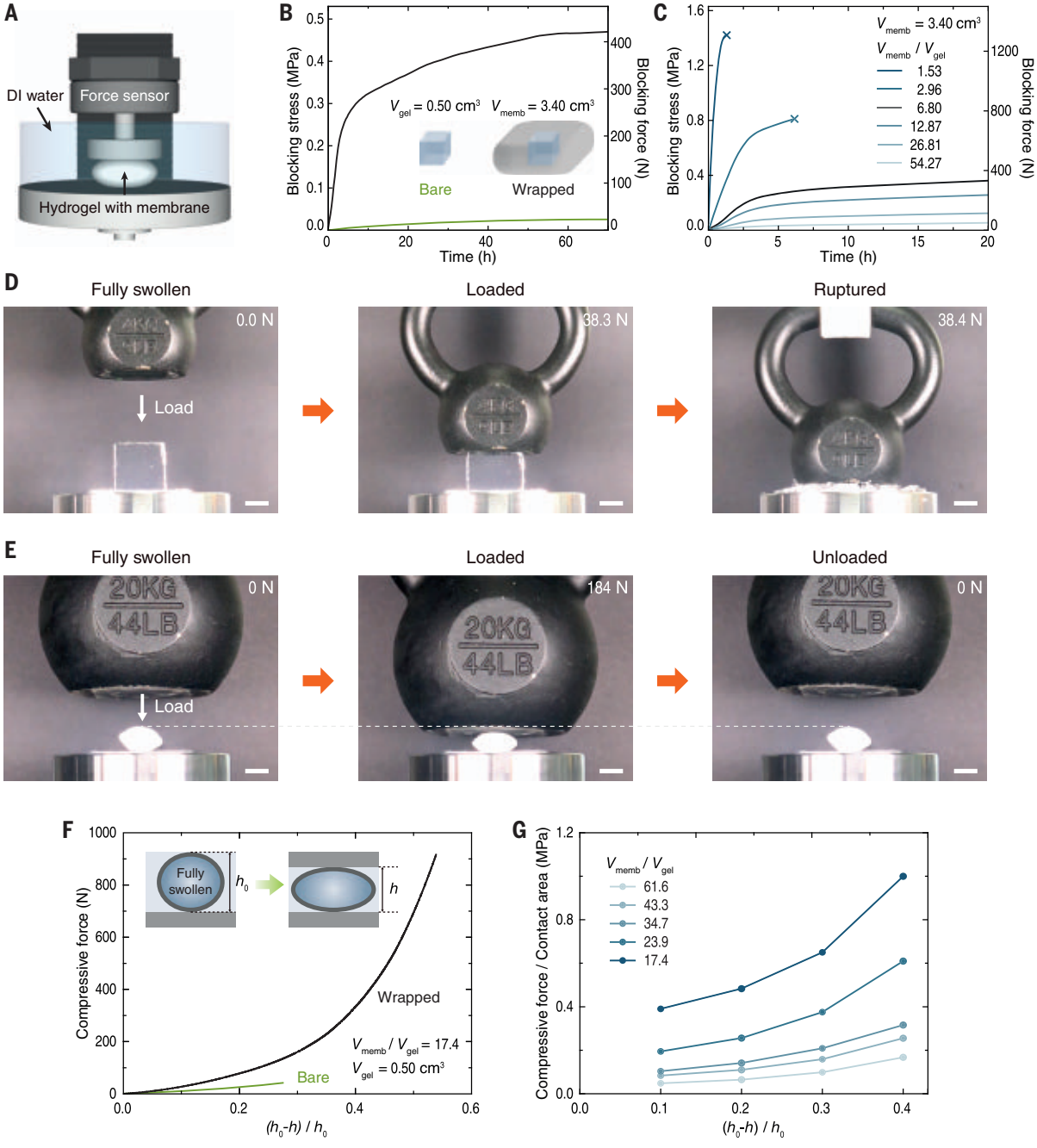


Fig. 2. Mechanical properties of the osmosis-driven hydrogel turgor actuator. (**A**) Schematic illustration of the setup for the constrained swelling experiment. A PSPA hydrogel wrapped in a selectively permeable membrane was placed between two rigid plates with a fixed height. (**B** and **C**) The stress generated by the turgor actuator was measured as a function of time. (B) Wrapped turgor actuator compared with a bare hydrogel. (C) Blocking stresses of the turgor actuators with different membrane-to-gel volumetric ratios were investigated. (**D** and **E**) Compression tests of a fully swollen bare hydrogel (D) and a turgor actuator (E). Weights were suspended from a force sensor and slowly lowered at 5 mm/s; 4- and 20-kg kettlebells were used, respectively. The bare gel ruptured at 38.3 N, whereas the wrapped gel endured 184 N and recovered its original shape when unloaded. Scale bars, 2 cm. (**F**) Compressive force-strain [$(h_0–h)/h_0$] curves for a bare hydrogel and a wrapped hydrogel with a swelling ratio of 17.4. (**G**) Compressive true stress, (compressive force)/(contact area), versus strain, $(h_0–h)/h_0$, curve for turgor actuators with the swelling ratio of the gel inside the membrane at various strains. The stiffness (tangential slope of compressive stress-strain curve) decreases with the swelling ratio (V_{memb}/V_{gel}), which is inversely related to the osmotic pressure of the turgor actuators.

the turgor actuator were measured as illustrated in Fig. 2A. Poly(3-sulfopropylacrylate) potassium salt (PSPA) polyelectrolyte hydrogels were used for the actuators because the ionizable polymer chains contribute to high osmotic pressure (see methods and fig. S3 for detailed fabrication of the turgor actuators) (*22*). As shown in Fig. 2B, the blocking stress of the turgor actuator reaches 0.46 MPa (421 N)—16.8 times as large as that of the bare hydrogel. The actuation speeds of the bare hydrogel and the turgor actuator are both very slow, taking about 4 days to reach the equilibrium swollen state (fig. S4).

We also measured the blocking stresses of the hydrogel turgor actuators with various

swelling ratios by changing the initial volume of the hydrogel (V_{gel}) inside the membrane (Fig. 2C). As the inner volume (V_{memb}) of the turgor actuators was the same for each, the gels with larger initial volume swelled less, retaining the higher osmotic pressures and exerting the large blocking stresses. For the same reason, the blocking stress of the actuator increases as the stroke decreases (fig. S5). The maximum stress of the hydrogel turgor actuator is 1.44 MPa (1330 N) with a swelling ratio of 1.53. The maximum stress was obtained immediately before the gel was squeezed out through the pores in the membrane in response to the excessively large turgor pressure developed inside the membrane (fig. S6).

The strength of the turgor actuator was also substantially higher than that of the bare hydrogel. To demonstrate the strength of the turgor actuators intuitively, dead weights were loaded on a bare hydrogel and a turgor actuator with the same amount of hydrogel (movie S1 and Fig. 2, D and E). Both were fully swollen in deionized water. The bare gel, which had a swelling ratio of 125 (fig. S7 and S8), was completely crushed under 38.4 N of load (Fig. 2D). By contrast, the turgor actuator containing the same hydrogel (swelling ratio of 17.4) in its membrane endured a much heavier weight (184 N) and recovered to its original state when unloaded (Fig. 2E). To quantitatively compare the two actuators in terms of strength, their compressive force-strain curves were measured. As shown in Fig. 2F, a hydrogel turgor actuator endured 917 N of compressive force before the membrane ruptured, which is a value 22 times as large as that of the fully swollen bare hydrogel (41 N). The compressive strength of the turgor actuators was limited by the durability of the wrapping, such as the adhesion strength or the toughness of the membrane.

To evaluate the effect of the osmotic pressure on the mechanical properties of the turgor actuators, true compressive stress-strain curves were measured at various swelling ratios. Strain is defined as $[(h-h_0)/h_0]$, where h_0 and h are the heights of the uncompressed and compressed samples, respectively. As shown in Fig. 2G, the true stress at the same strain increases with the osmotic pressure of the hydrogel inside the membrane. Further, the stiffness of the turgor actuator (tangential slope of the compressive stress-strain curve) increases with the osmotic pressure (see supplementary text). Altogether, the mechanical properties of a turgor actuator can be manipulated by controlling the swelling ratio of the hydrogel.

Although hydrogel turgor actuators can generate large actuation stress by using the osmotic pressure, actuation speeds are fairly low because the actuations are driven by the diffusion of water. To achieve fast actuation, we added an electroosmotic effect to that of the osmosis. Electroosmosis is a constant and rapid water flow through the electric double layer (EDL) of charged porous materials under an electric field. As illustrated in Fig. 3A, a polyelectrolyte hydrogel contains fixed charges bound to the polymer network, forming an EDL (*23*). Thus, the counterions in the electrolyte solution can migrate across the charged polymer mesh under an electric field, thereby dragging the water into the gel network (*24*). The dragged water molecules are captured by the hydrophilic polymer chain, which leads to the swelling of the hydrogel. This active transport of water allows the polyelectrolyte gels to swell much faster than they do in osmosis-driven swelling.

To characterize the electroosmosis-driven actuation of the hydrogel turgor actuator, we measured the blocking stress and speed. As shown in Fig. 3B, the electroosmotic actuation (V_{memb}/V_{gel} = 2.94, E = 12 V/cm) was much faster than the osmosis-driven actuation. The electroosmotic stress generation rate (~0.23 MPa/min) was about 22 times as fast as that of the corresponding osmotic rate (~0.01 MPa/min) at the initial stage of actuation (30 s). The maximum stress generated by the electroosmotic actuation (~0.79 MPa) is similar to that generated by the osmotic actuation (~0.73 MPa) with the gel squeezed out through the membrane by the high turgor pressure. However, without squeeze-out of the gel, the actuation stress and speed increased with the electric field intensity (Fig. 3F and table S2). The improvement in actuation stress by electroosmosis was also observed in various strokes (fig. S5). In a stress relaxation test, the maximum and relaxed stresses of the electroosmotic turgor actuator are both much higher than those of the osmotic turgor actuator (fig. S10 and supplementary text). The effects of the cross-linker density, the sign of the charged monomer, and the external solution concentration on the electroosmotic actuation were investigated (fig. S11 and supplementary text). As demonstrated in Fig. 3C, this force generator can break a rigid brick. Under a 4-V/cm electric field, the turgor actuator broke a brick within 5 min (movie S2). The critical stress used to break the brick was ~0.38 MPa (570 N) (fig. S12). Further, a turgor actuator composed of a hydrogel with a volume of 4.3 cm^3 lifted a 21-kg weight to twice the initial height (fig. S13). From the area under the force-stroke curve (fig. S5) and the slope of the stroke-time curves (fig. S14), the work density and the power density of the turgor actuator were respectively calculated as 7.0 MJ/m^3 and 2.33 kW/m^3, thereby outperforming current hydrogels (~10^{-2} kJ/m^3 and ~10^{-5} kW/m^3) (*17*). Additionally, reversible actuation of the turgor actuator can be achieved by controlling the electric field (fig. S15 and supplementary text).

We theoretically analyzed the kinetics of the electroosmotic actuation of the turgor actuator. Electroresponsive hydrogels have been conventionally used as bending actuators, in which the bending direction was the primary interest. Most previous studies have focused on the interface between the fully swollen hydrogel and the surrounding solution (*9*, *25*, *26*). The rate of electroosmotic flow (EOF) in a porous medium (Q_{in}), the pore size of which is much larger than the Debye length, is given by (*27*)

$$Q_{in} = \frac{\varepsilon E \zeta \phi A}{\mu} \quad (7)$$

where ε, E, ζ, ϕ, μ, and A are the permittivity of the liquid, an electric field, the zeta potential of the medium, the porosity of the medium, the viscosity of the liquid, and the inlet area, respectively. The electroosmotic inflow (Q_{in}) is either absorbed by the hydrogel (Q_{abs}) or passes through the hydrogel (Q_{out}): $Q_{in} = Q_{abs} + Q_{out}$ (Fig. 3E, inset). In the early stages of the electroosmosis-driven swelling, the entire electroosmotic inflow is absorbed by the gel: $Q_{in} \simeq Q_{abs}$. This is because water dragged into the hydrogel from the solution by electroosmosis is captured by the hydrophilic polymer network, as shown in fig. S16. When the hydrogel reaches the swelling equilibrium, the EOF merely passes through the hydrogel: $Q_{abs} \simeq 0$.

As shown in Fig. 3D, the mass of the gel increases linearly over time at the initial stage, the slopes of which increase with the electric field intensity. Swelling ceases when a gel is fully swollen. Under an electric field of 1 to 6 V/cm, the average swelling rate until reaching the fully swollen state is 43.5 to 205 times as fast as that of pure osmosis. Fig. 3E shows that the measured swelling rates are consistent with our theory that predicts the linear increase of the swelling rate with the intensity of the electric field (see methods for details). By measuring flow rates through saturated hydrogels, we found that the outlet flow rates (Q_{out}) also increase linearly with the electric field intensity (fig. S17). We also found that the mass of the fully swollen gel by an electric field >2 V/cm is always 150% as large as that by osmosis (fig. S18). This is attributed to the fact that the equilibrium volume of a hydrogel is associated with the stretch limit of a polymer network (*28*).

We model the actuation stress exerted by the turgor actuators. Under an electric field, a polyelectrolyte gel can absorb the water by both osmosis and electroosmosis. Under osmosis, the osmotic pressure (π_{os}) induces the diffusion of the liquid into the gel, whereas the membrane pressure resulting from osmosis ($P_{memb,os}$) squeezes out the liquid. The diffusion flux associated with osmosis (q_{os}) follows Fick's law $q_{os} = -DcV_m|\nabla P|/(RT)$ where D, c, V_m, P, R, and T respectively denote the self-diffusion coefficient of the liquid, the liquid concentration of the polyelectrolyte gel, the molar volume of the liquid, the pressure, the gas constant, and the temperature (*29*). The

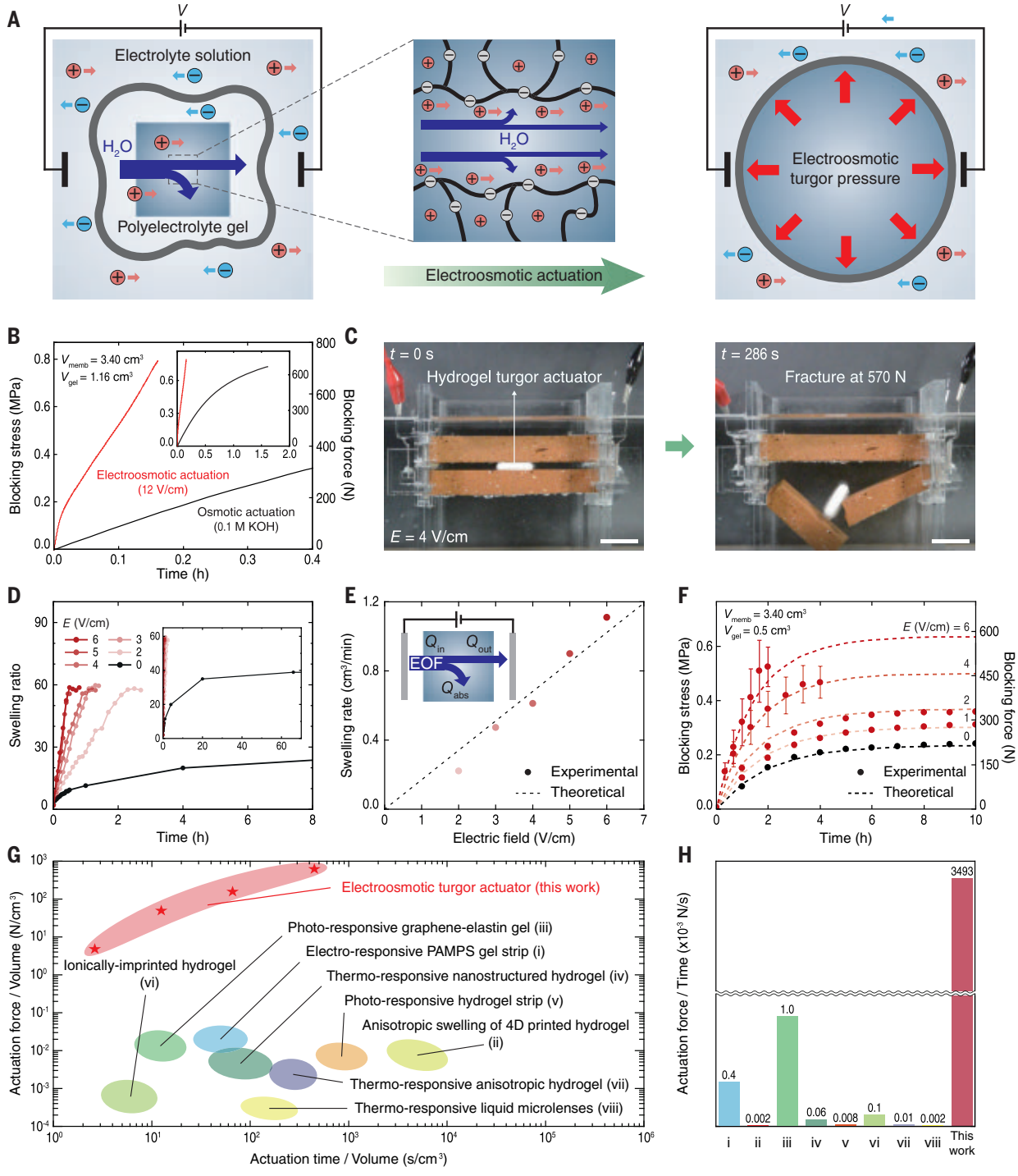


Fig. 3. Electroosmotic actuation of the hydrogel turgor actuator in electrolyte solution. (**A**) Schematic illustration of the electroosmotic actuation process of a hydrogel turgor actuator in an electrolyte solution. The turgor actuator was placed between platinum electrodes in KOH solution. When a voltage was applied, potassium ions migrated through the negatively charged polymer mesh, causing electroosmotic flow (EOF) inside the gel. The flow caused the gel to swell rapidly, generating a large turgor pressure inside the membrane. (**B**) Blocking stress versus time curves for the hydrogel turgor actuator actuated by osmosis and electroosmosis at zero stroke. The same turgor actuators (V_{memb} = 3.40 cm^3, V_{gel} = 1.16 cm^3) were used for each measurement. The inset shows the stress versus time over a long time scale. (**C**) Brick breaking within 5 min with turgor actuator operated by electroosmosis with a 4-V/cm electric field. The average stress at the fracture of the brick was ~0.38 MPa (570 N). Scale bar, 3 cm. (**D**) Swelling ratio change of bare hydrogels under different electric field intensities. (**E**) Electric field dependence of the swelling rate of bare gels with an electroosmotic actuation where $Q_{in} \simeq Q_{abs}$. The gel absorbed water at the rate of Q_{abs} as a part of the water inflow rate Q_{in}, whereas the remainder (Q_{out}) flowed out of the gel. (**F**) Blocking stress of turgor actuator (V_{memb} = 3.40 cm^3, V_{gel} = 0.5 cm^3) versus time with different electric field intensities. The dashed line shows the theoretical prediction of the stress evolution of hydrogel turgor actuators. Error bars denote SDs; N = 3. (**G**) Ashby plot of the actuation force and time, which are normalized by volume of the hydrogel, for the electroosmotic turgor actuators and other osmotic actuators. The hydrogel turgor actuators exhibited the largest force and the fastest speed simultaneously. The data used are summarized in table S2. (**H**) Force generation rates (that is, actuation force divided by corresponding actuation time) of electroosmotic turgor actuators and other osmotic actuators.

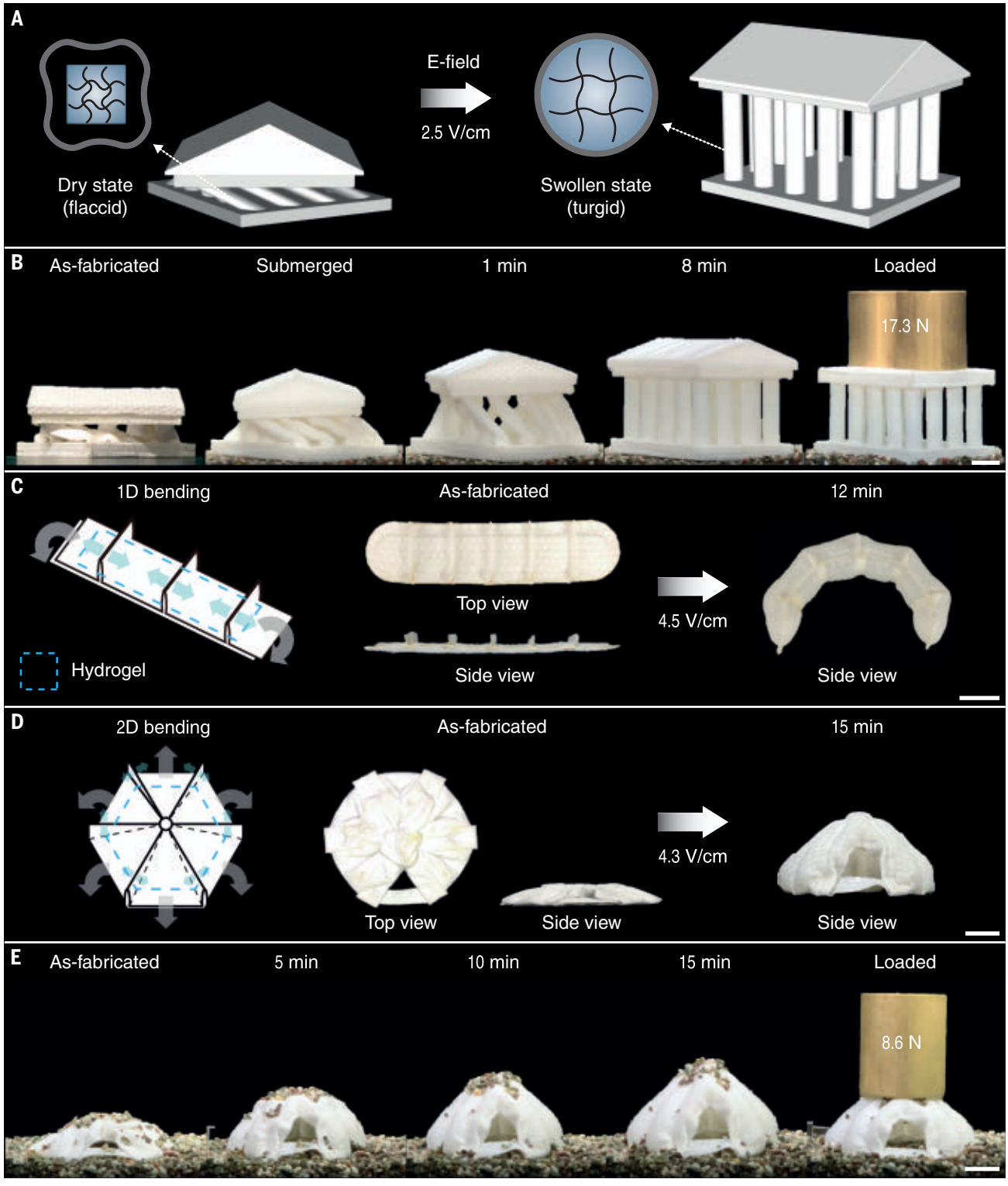


Fig. 4. Rapid construction of underwater structures. (**A**) Schematic illustrations of the construction of a Greek temple by electroosmotic actuation. The pillars of the Greek temple were initially filled with dry hydrogel and thus started from a fully collapsed structure. Actuation was conducted at 2.5 V/cm in 0.1-M KOH aqueous solution. (**B**) Rapid construction of the temple with actuation time ~8 min. The completed temple endured two weights measuring 19.6 N, and their effective load is calculated to be 17.3 N by considering the buoyancy. (**C**) Schematic illustration and pictures of the 1D bending structure with joints. When the hydrogel swelled, the wrinkles unfolded in the direction of the sky-blue arrow and the system bent in the direction of the gray arrow. The actuation was conducted under a 4.5-V/cm electric field for 12 min. (**D**) Schematic illustration and pictures of 2D bending resulting in the formation of an igloo. The actuation was conducted under a 4.3-V/cm electric field. (**E**) Actuation progress for construction of the igloo for 15 min. The complete igloo endured the effective load (8.6 N) considering buoyancy. Scale bars, 2 cm.

pressure driving the diffusion is given by $\Delta P = \pi_{os} - P_{memb,os}$, and the diffusion flow rate is written as $Q_{os} = DcV_m A\Delta P/(\delta RT)$, where A is the liquid inlet area and δ is the ion concentration boundary layer thickness. For the turgor actuator, c and δ are assumed constant because the gel volume is fixed in the relatively stiff membrane. The membrane pressure corresponds to the mechanical stress developed inside the gel that is restrained from swelling, $P_{memb,os} = K\gamma_o$, with K and γ_o respectively being the bulk modulus of the polyelectrolyte gel and the restrained volumetric strain due to osmosis (*30*). Because γ_o originated from $\int Q_{os}dt/V_0$, we write $P_{memb,os} = K\int Q_{os}dt/V_0$ with V_0 being the volume of the polyelectrolyte gel turgor actuator. By substituting $Q_{os} = V_0\left(dP_{memb,os}/dt\right)/K$ in Fick's law above and solving, we get

$$P_{memb,os} = \pi_{os}\left(1 - e^{-t/\tau_o}\right) \quad (8)$$

where $\tau_o = RTV_0\delta/(DcV_mKA)$.

Under electroosmosis, the electroosmotic pressure drives the liquid flow into the polyelectrolyte gel, whereas the membrane pressure resulting from electroosmosis ($P_{memb,eos}$) tends to drive the liquid flow out of the gel. The inflow driven by electroosmosis is expressed as Eq. 7. The outflow driven by the membrane pressure follows Darcy's law, $q_e = -(k/\mu)|\nabla P_{memb,eos}|$, which describes the viscous liquid flux q_e through a porous medium with permeability k. Therefore, the total inlet flow rate is the difference between the electroosmotic inflow rate and the membrane pressure-driven outflow rate: $Q_e = \left(\varepsilon\zeta E\phi - k\Delta P_{memb,eos}/L\right)A/\mu$, where L is the characteristic length of the polyelectrolyte gel. Q_e causes membrane pressure as a result of a volumetric strain $P_{memb,eos} = K\int Q_e dt/V_0$. Using $Q_e = V_0\left(dP_{memb,eos}/dt\right)/K$, we get

$$P_{memb,eos} = \eta_e\left(1 - e^{-t/\tau_e}\right) \quad (9)$$

where $\eta_e = \varepsilon\zeta E\phi L/k$ and $\tau_e = \mu LV_0/(kKA)$. Therefore, the total membrane pressure $P_{memb} = P_{memb,os} + P_{memb,eos}$ is obtained as

$$P_{memb} = \pi_{os}\left(1 - e^{-t/\tau_o}\right) + \eta_e\left(1 - e^{-t/\tau_e}\right) \quad (10)$$

meaning, as shown in Eqs. 5 and 6, that blocking stress equals the water inflow pressure in the equilibrium state: $\sigma_{block}^{wrapped} \simeq P_{memb} \simeq \pi_{os} + \eta_e = P_{in}$. Finally, the total membrane stress generates the actuation force of the hydrogel turgor actuator as

$$F = \sigma_{block}^{wrapped}A_c \simeq P_{memb}A_c = \pi_{os}A_c\left(1 - e^{-t/\tau_o}\right) + \eta_e A_c\left(1 - e^{-t/\tau_e}\right) \quad (11)$$

where A_c is the area over which the force is applied.

Using the parameter values obtained from the literature and experiments as described in table S4 and the supplementary text, we calculated the theoretical actuation stress of the turgor actuator and plotted the results (dashed lines) in Fig. 3F. We can see that the calculated actuation stresses are consistent with the experimental data under electric fields ranging from 0 to 6 V/cm. We can also see that the generated blocking stress approached the osmotic pressure of the turgor actuator once the electric field was turned off (fig. S20).

Overall, the electroosmotic turgor actuators exhibited stronger and faster actuations compared to typical osmotic hydrogel actuators (table S5), especially with actuation force increased by a factor of 10^2 to 10^6 (Fig. 3G). The acceleration in swelling by electroosmosis (Fig. 3D), combined with the design employing turgor pressure, leads to substantially faster force generation (3.5 N/s) than that of the osmotic actuator (below 0.001 N/s) (Fig. 3H).

We used the hydrogel turgor actuators as a structural material in an aqueous environment because they can swell rapidly and withstand large forces (movies S3 and S4). We first demonstrated the rapid construction of a Greek temple structure, the columns of which were made of a membrane filled with a small amount of polyelectrolyte gel (Fig. 4A). As shown in Fig. 4B, the columns were flaccid before applying the electric field, so the roof and the floor of the temple were in contact. However, when the electric field of 2.5 V/cm was applied, the columns gradually lifted the roof for ~8 min, thus forming a complete Greek temple structure. This underwater temple endured a 17.3 N of effective load, considering buoyancy.

We further demonstrated the construction of a complicated structure by incorporating wrinkles within the membrane. The wrinkles on the membrane act as joints which unfold when the gel swells inside the membrane. A rod-shaped 1D actuator showed in-plane bending (Fig. 4C), whereas a planar 2D actuator showed out-of-plane bending (Fig. 4D). Two-dimensional bending of the planar actuator led to creation of an igloo-like shelter structure within 15 min of application of a 4.3-V/cm electric field. The underwater igloo endured a weight of 8.6 N, considering buoyancy. Even without an electric field, a turgor actuator can still serve as a support as a result of osmotic pressure. Thus, constructed structures can continually maintain their shape with high stiffness in an aqueous environment.

We proposed a hydrogel-based strong and fast turgor actuator that converts high osmotic pressure into a corresponding strong actuation stress with the aid of a membrane. Electroosmotic actuation makes the turgor actuator substantially faster and stronger by the active transport of water into the hydrogel. The actuators were used to break a rigid brick and build complex underwater structures within a few minutes. Our dynamic model of the actuation is expected to serve as a guideline for the control of a prospective soft robot. Our strategies can be applied to other elastomer- or hydrogel-based actuators, as they provide enhanced mechanical power while retaining the intrinsic advantages of the materials and thereby extend the scope of applications.

REFERENCES AND NOTES

1. Y. Lee *et al.*, *Sci. Robot.* **5**, eaaz5405 (2020).
2. I. Must, E. Sinibaldi, B. Mazzolai, *Nat. Commun.* **10**, 344 (2019).
3. M. Choi *et al.*, *Nat. Photonics* **7**, 987–994 (2013).
4. D. J. Beebe *et al.*, *Nature* **404**, 588–590 (2000).
5. S. Zhang *et al.*, *Nat. Mater.* **14**, 1065–1071 (2015).
6. H. Qin, T. Zhang, N. Li, H. P. Cong, S. H. Yu, *Nat. Commun.* **10**, 2202 (2019).
7. J. Zheng *et al.*, *J. Mater. Chem. C Mater. Opt. Electron. Devices* **6**, 1320–1327 (2018).
8. Z. Han *et al.*, *ACS Appl. Mater. Interfaces* **12**, 12010–12017 (2020).
9. D. Morales, E. Palleau, M. D. Dickey, O. D. Velev, *Soft Matter* **10**, 1337–1348 (2014).
10. S.-L. Xiang, Y.-X. Su, H. Yin, C. Li, M.-Q. Zhu, *Nano Energy* **85**, 105965 (2021).
11. S. Juodkazis *et al.*, *Nature* **408**, 178–181 (2000).
12. L. W. Xia *et al.*, *Nat. Commun.* **4**, 2226 (2013).
13. Z. Jiang, B. Diggle, I. C. G. Shackleford, L. A. Connal, *Adv. Mater.* **31**, e1904956 (2019).
14. Z. Jiang *et al.*, *Chem. Mater.* **33**, 7818–7828 (2021).
15. P. Song, Y. F. Zhang, J. Z. Kuang, *J. Mater. Sci.* **42**, 2775–2781 (2007).
16. Y. S. Kim *et al.*, *Nat. Mater.* **14**, 1002–1007 (2015).
17. Y. Ma *et al.*, *Sci. Adv.* **6**, eabd2520 (2020).
18. H. Yuk *et al.*, *Nat. Commun.* **8**, 14230 (2017).
19. X. Liu, J. Liu, S. Lin, X. Zhao, *Mater. Today* **36**, 102–124 (2020).
20. J. Li, Y. Hu, J. J. Vlassak, Z. Suo, *Soft Matter* **8**, 8121–8128 (2012).
21. W. R. K. Illeperuma, J.-Y. Sun, Z. Suo, J. J. Vlassak, *Soft Matter* **9**, 8504–8511 (2013).
22. J. Li, Z. Suo, J. J. Vlassak, *Soft Matter* **10**, 2582–2590 (2014).
23. X. Zhou, K. Zhao, *Phys. Chem. Chem. Phys.* **19**, 20559–20572 (2017).
24. J. Ko *et al.*, *ACS Nano* **14**, 11906–11918 (2020).
25. M. Doi, M. Matsumoto, Y. Hirose, *Macromolecules* **25**, 5504–5511 (1992).
26. T. Shiga, Y. Hirose, A. Okada, T. Kurauchi, *J. Appl. Polym. Sci.* **46**, 635–640 (1992).
27. R. F. Probstein, "Electroosmosis" in *Physicochemical Hydrodynamics: An Introduction* (Wiley, ed. 2, 1994), pp. 197–202.
28. M. Rubinstein, R. H. Colby, in *Polymer Physics*. (Oxford Univ. Press, 2003), pp. 74–78.
29. R. B. Bird, W. E. Stewart, E. N. Lightfoot, *Transport Phenomena* (Wiley, 2006).
30. G. Lin, S. Chang, C.-H. Kuo, J. Magda, F. Solzbacher, *Sens. Actuators B Chem.* **136**, 186–195 (2009).

ACKNOWLEDGMENTS

Funding: This work was supported by National Research Foundation of Korea (NRF) grants funded by the Korean Government (2018M3A7B4089670, 2018-052541, and 2021R1A2C2092737). The Institute of Engineering Research at Seoul National University provided the research facilities for this work. **Author contributions:** H.-Y.K., and J.-Y.S. supervised the research. All authors contributed to interpreting the results and preparing the manuscripts. H.N., Y.-W.K., and C.S.P. conceived the concept, designed the experiments, fabricated the devices, and collected data. S.J. developed the theoretical modeling and performed data analysis. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm7862
Materials and Methods
Supplementary Text
Figs. S1 to S26
Tables S1 and S5
References (*31*–*44*)
Movies S1 to S4

12 October 2021; accepted 16 March 2022
10.1126/science.abm7862

3D PRINTING

Volumetric additive manufacturing of silica glass with microscale computed axial lithography

Joseph T. Toombs[1]*, Manuel Luitz[2], Caitlyn C. Cook[3], Sophie Jenne[2], Chi Chung Li[1], Bastian E. Rapp[2,4,5,6], Frederik Kotz-Helmer[2,4,5], Hayden K. Taylor[1]*

Glass is increasingly desired as a material for manufacturing complex microscopic geometries, from the micro-optics in compact consumer products to microfluidic systems for chemical synthesis and biological analyses. As the size, geometric, surface roughness, and mechanical strength requirements of glass evolve, conventional processing methods are challenged. We introduce microscale computed axial lithography (micro-CAL) of fused silica components, by tomographically illuminating a photopolymer-silica nanocomposite that is then sintered. We fabricated three-dimensional microfluidics with internal diameters of 150 micrometers, free-form micro-optical elements with a surface roughness of 6 nanometers, and complex high-strength trusses and lattice structures with minimum feature sizes of 50 micrometers. As a high-speed, layer-free digital light manufacturing process, micro-CAL can process nanocomposites with high solids content and high geometric freedom, enabling new device structures and applications.

The uses of glass are innumerable because of its optical transparency, thermal and chemical resistance, and low coefficient of thermal expansion. Established applications in architecture, consumer products, optical systems, and art have been joined by specialized uses such as fiber optics in communication, diffractive optics in augmented reality, and lab-on-a-chip devices for chemical and biological analyses (*1*–*3*). With increased specialization come more demanding requirements for geometry, size, and optical and mechanical properties. Additive manufacturing (AM) has emerged as a promising technique to meet challenging new combinations of requirements. AM of glass materials has been achieved with fused filament fabrication of molten glass (*4*, *5*), selective laser melting of pure glass powder (*6*, *7*), direct ink writing of silica sol-gel inks (*8*), stereolithography (SLA) (*9*–*11*), and multiphoton direct laser writing (DLW) (*12*) of silica nanocomposites that consist of silica nanoparticles dispersed in a photopolymerizable organic liquid.

All these methods use serial material deposition or conversion, which can limit geometric freedom. Layering-induced defects can also affect the printed object's optical and mechanical properties (*5*, *13*). We introduce volumetric AM (VAM) of glass nanocomposites. VAM describes techniques that polymerize whole three-dimensional (3D) objects simultaneously in a volume of precursor material, circumventing the need to build objects layer by layer. VAM methods based on holographic exposure (*14*), orthogonal superposition (*15*), and tomographic principles (*16*, *17*) are enabled by specialized optical engineering and photopolymer synthesis. The tomographic technique of computed axial lithography (CAL) polymerizes 3D structures by the azimuthal superposition of iteratively optimized light projections from temporally multiplexed exposures (Fig. 1A) (*16*, *18*, *19*). CAL has several advantages for processing glass nanocomposites. No relative motion occurs between the precursor material and the fabricated object during printing, so high-viscosity and thixotropic nanocomposite precursors can easily be used. The layerless nature of the process enables smooth surfaces and complex geometries. Because the fabricated object is surrounded by precursor material during printing, sacrificial solid supporting structures are not needed. These attributes are desirable for applications that include micro-optical components and microfluidics.

We sought the production of microscale features, so we constructed a "micro-CAL" apparatus (Fig. 1B) that coupled a laser light source into an optical fiber with small mode field size and low numerical aperture (*17*) and demagnified the light pattern defined by the digital micromirror device. This design minimized the system's étendue and hence the divergence and blurring of light. We measured optical resolution in terms of the modulation transfer function (MTF), which is the level of contrast transfer by the complete optical system as a function of spatial frequency. We achieved an MTF greater than 0.4 at frequencies ≥66.7 cycles mm^{-1} in the central 1.5-mm diameter of the build volume (Fig. 1, D and E, and figs. S2 to S4). Combined with gradient descent digital mask optimization (*16*), the micro-CAL system enabled rapid printing (within about 30 to 90 s) of microstructures with minimum feature sizes of 20 and 50 μm in polymer and fused silica glass, respectively.

For the fused silica prints, we used a photocurable micro-stereolithography (μSL) v2.0 nanocomposite with high transparency (Fig. 2A) consisting of a liquid monomeric photocurable binder matrix and 35 vol % solid amorphous spherical silica nanoparticles with a nominal diameter of 40 nm. The high–solids content nanocomposite had a zero shear viscosity of 10 Pa·s at 23°C, and it exhibited thixotropic shear-thinning properties at moderate shear rates (1 to 100 s^{-1}) and shear-thickening properties at high shear rates (>100 s^{-1}) (*20*). The binder was polymerized via free-radical polymerization and supported the nanoparticles in the printed construct. After printing, we removed structures from the volume of nanocomposite and reused surplus nanocomposite for later prints. We developed the structures by rinsing in ethanol or propylene glycol methyl ether acetate for about 10 min to remove excess uncured nanocomposite. Heating up to 60°C was applied to reduce viscosity by up to an order of magnitude to assist in the development of small features. We subjected the resulting green parts to thermal treatment in two steps: debinding and sintering (Fig. 1A and tables S3 and S4). The debinding treatment burned out the polymer binder matrix, resulting in a porous silica brown part. During sintering, the nanoparticles of the brown part fused together, forming a dense transparent glass part. An isotropic linear shrinkage (fig. S8) of 26% occurred during sintering, which was consistent with the theoretical shrinkage predicted by thermogravimetric analysis, so it was necessary for us to scale parts in computer-aided design before fabrication to account for the dimensional change (*20*).

The tomographic illumination process of CAL means that material that is outside the target geometry receives an appreciable light dose. To achieve selective material conversion, the resin precursor therefore has a threshold light exposure dose below which polymerization is negligible. In prior VAM research, the induction period—the period of time in which conversion is inhibited by radical scavenger species in the resin—was a result of oxygen inhibition (*15*–*17*). However, the glass nanocomposite used in this work exhibited a small natural induction period. Besides molecular oxygen, several molecules—including quinones and nitroxides, such as 2,2,6,6-tetramethylpiperidinoxyl (TEMPO)—are recognized as effective radical inhibitors (*21*, *22*). We added various concentrations of TEMPO to the nanocomposite and performed real-time ultraviolet (UV) Fourier transform infrared spectroscopy (FTIR) analysis to determine the effect of TEMPO concentration on the inhibition time (Fig. 2B). The addition of TEMPO increased the duration of the induction period

[1]Department of Mechanical Engineering, University of California, Berkeley, CA 94720, USA. [2]Department of Microsystems Engineering, Albert Ludwig University of Freiburg, 79104 Freiburg, Germany. [3]Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. [4]Glassomer GmbH, Georges-Köhler-Allee 103, 79110 Freiburg, Germany. [5]Freiburg Materials Research Center (FMF), Albert Ludwig University of Freiburg, 79104 Freiburg, Germany. [6]Freiburg Center of Interactive Materials and Bioinspired Technologies (FIT), Albert Ludwig University of Freiburg, 79110 Freiburg, Germany.
*Corresponding author. Email: jtoombs@berkeley.edu (J.T.T.); hkt@berkeley.edu (H.K.T.)

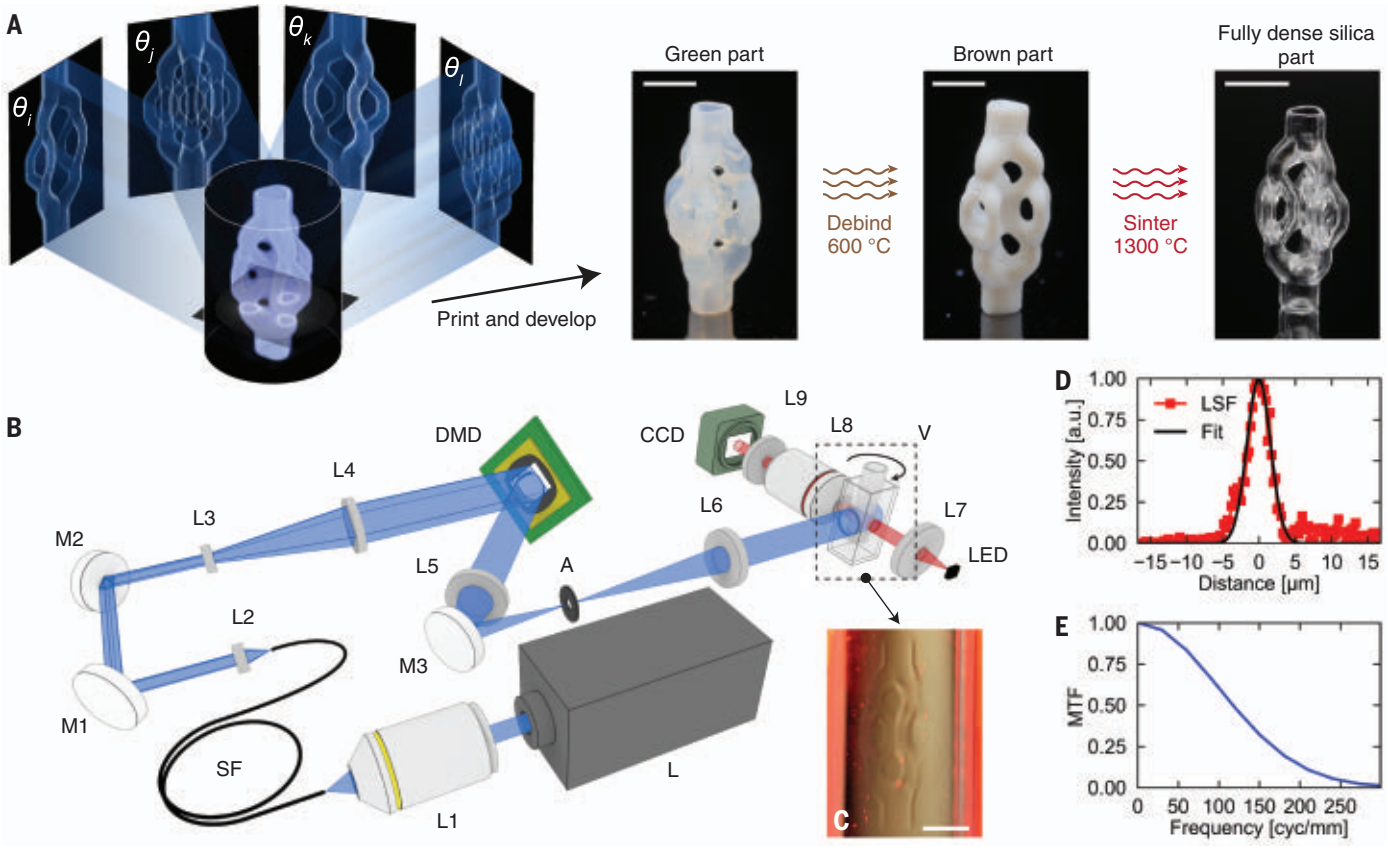


Fig. 1. Printing transparent fused silica glass with micro-CAL. (**A**) Tomographic superposition produces a 3D light dose that selectively polymerizes a geometry. After printing, the part is developed by rinsing away residual nanocomposite in a solvent. Debinding and sintering steps follow. Scale bars are 2 mm. θ_i to θ_l, examples of optimized tomographic digital light projections. (**B**) Optomechanical setup. A detailed description of the setup and components are given in (*20*). A, aperture; CCD, charge-coupled device; DMD, digital micromirror device; L, 3-W, 442-nm laser diode; L1 to L9, lenses 1 to 9; LED, light-emitting diode; M1 to M3, mirrors 1 to 3; SF, multimode square core fiber; V, print container. (**C**) Immediately after light exposure, the printed object can be observed in the container, weakly distorting the background. Scale bar is 2.5 mm. (**D** and **E**) Line spread function (LSF) (D) and MTF (E) at the focal plane of the micro-CAL system. a.u., arbitrary units; cyc/mm, cycles per millimeter.

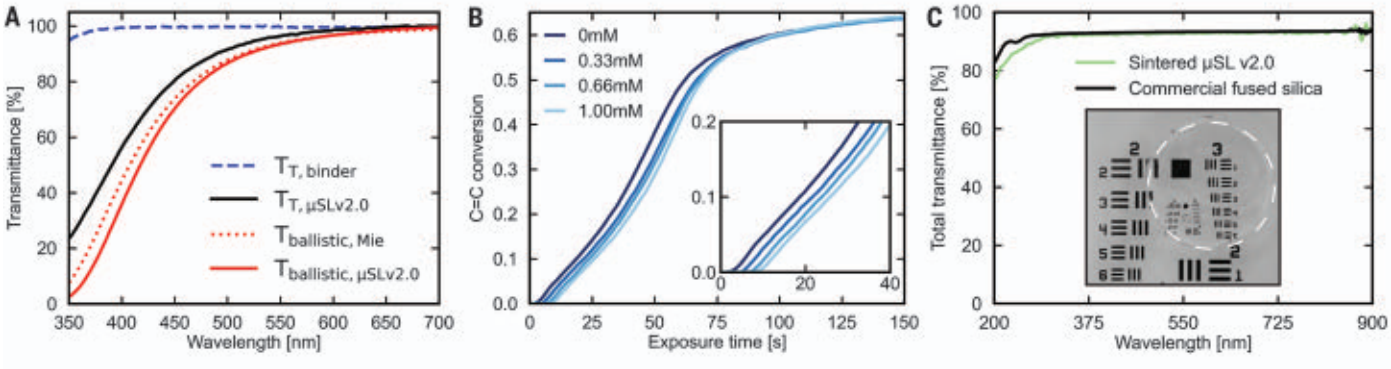
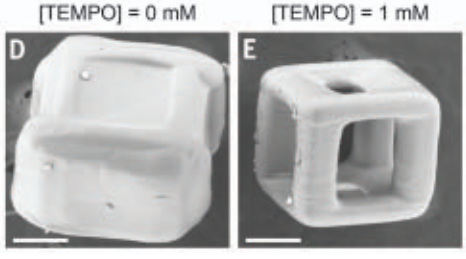


Fig. 2. Nanocomposite and sintered silica material characterization. (**A**) Binder (nanocomposite without silica) $T_{T,binder}$ and μSL v2.0 (with silica) transmittance spectra, where $T_{T,\ \mu SLv2.0}$ is total transmittance, $T_{ballistic,\ \mu SLv2.0}$ is the ballistic (collimated) component, and $T_{ballistic,\ Mie}$ is the transmittance estimated by Mie theory. (**B**) UV FTIR measurement of carbon-carbon double-bond conversion as a function of exposure time and concentration of TEMPO. An increased induction period is observed with increased TEMPO concentration. (**C**) Total transmittance spectra of micro-CAL–printed and sintered μSL v2.0 disk and commercial fused silica coverslip. The inset image shows qualitatively the optical transparency of a printed circular disk (inside the dashed circle) in front of a US Air Force (USAF) resolution target. See (*20*) for experimental details. (**D** and **E**) Scanning electron microscope (SEM) micrographs of cube cages printed and developed with 0 and 1 mM TEMPO, respectively. Scale bars are 200 μm.



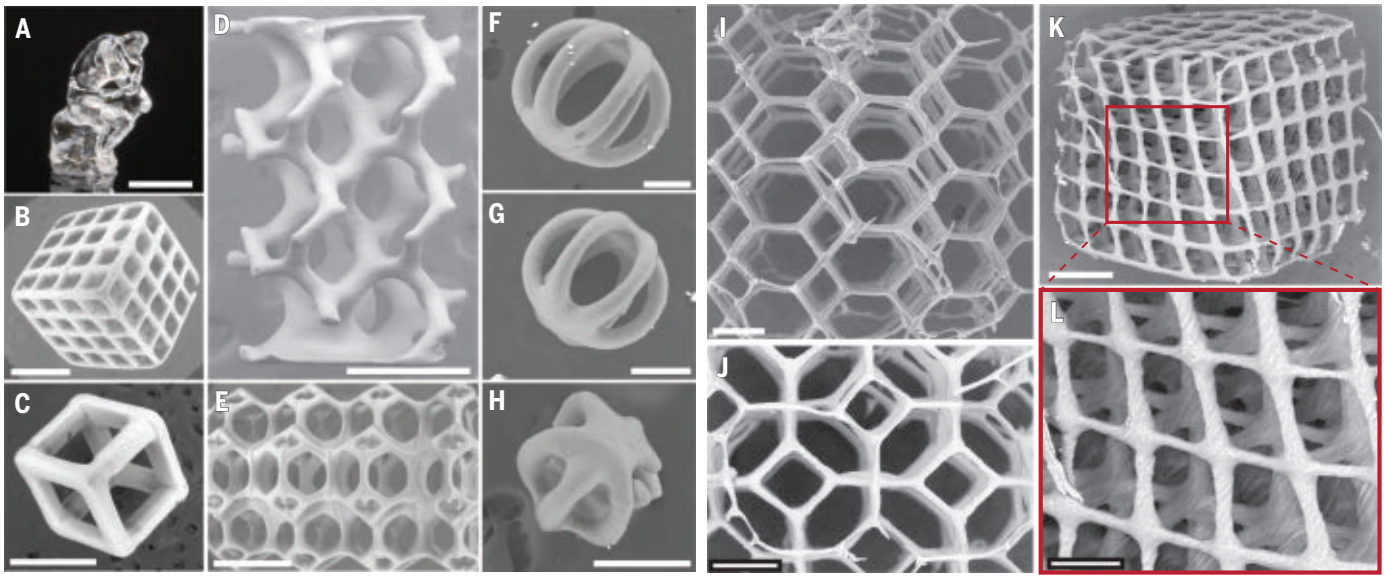


Fig. 3. CAL-printed glass structures. (**A** to **L**) Shown are Rodin's *The Thinker* (A), cubic lattice structures [(B) and (C)], a skeletal gyroid lattice with minimum positive feature size of 50 μm (D), a tetrakaidecahedron lattice (E), and spherical cage structures with minimum positive feature sizes of 75 (F), 60 (G), and 50 μm (H), respectively. Also shown are tetrakaidecahedron lattices [(I) and (J)] and a cubic lattice printed in monomeric photopolymer with a minimum positive feature size of 20 μm in each [(K) and (L)]. The images are a photograph in (A) and SEM micrographs in (B) to (L). Scale bars are 1 mm [(A) to (E)], 200 μm [(F) to (H)], 250 μm [(I) and (J)], 500 μm (K), and 100 μm (L).

and had a negligible effect on the kinetics of polymerization and the maximum degree of conversion.

The sharply nonlinear relationship between conversion and exposure dose provided by TEMPO substantially increased the lithographic contrast of the resin, thereby improving the feature modulation of the micro-CAL process. Increasing the TEMPO concentration reduced the degree of conversion in regions surrounding the printed object and inside internal voids and interstitial regions within the object (fig. S21). We show an example of such improvement using a particular set of digital light projections (Fig. 2, D and E). With TEMPO, conversion inside the void of the cubic cage was reduced, and the removal of uncured material was more easily achieved than in the nanocomposite without TEMPO (*20*). This improvement enabled fabrication of diverse geometries with positive feature sizes as low as 50 μm in the μSL v2.0 material (Fig. 3). In pure monomeric resin precursors, we achieved substantially smaller positive feature sizes as low as 20 μm (Fig. 3, I to L). We attribute this resolution enhancement for polymeric structures to the absence of solid nanoparticles, which results in easier development owing to lower resin precursor viscosity, a less brittle green state, and less light scattering (*20*).

Synthetic microstructured cellular materials have found use in a variety of fields, including photonics, energy, bioengineering, and desalination, as well as in high-temperature environments (*23*–*25*). Specifically, mechanical metamaterials, designed to exhibit mechanical properties that are unattainable by the bulk material—for example, a negative Poisson ratio—are emerging as an important area in AM because they often have a porous nature that is challenging to reproduce with conventional manufacturing techniques (*26*). In contrast with SLA, DLW, and fused filament fabrication, CAL builds objects volumetrically, which means that complex, low–relative density lattice and truss structures can be created in any orientation without supporting material (Figs. 3, B to D, and 4A). We fabricated tetrakaidecahedron lattices from transparent fused silica glass with strut elements of about 100 μm in diameter (Fig. 3E). For specific applications in which the orientation of the microstructure is critical, volumetric processing may prove useful because it eliminates defects due to layering that would be present in certain print orientations using other AM techniques.

We obtained surface-roughness metrics relevant to flaw-size characterization: arithmetic mean surface height, maximum valley depth, and root mean square surface gradient for rectangular beam specimens printed with micro-CAL and with SLA in vertical and horizontal orientations with respect to the build plate (fig. S13). We observed that micro-CAL produced significantly smaller and blunter defects and overall smoother surfaces (*20*). Three-point-bending mechanical testing showed that the difference between average fracture stress for different AM modalities was not statistically significant (Fig. 4A). However, the Weibull modulus for micro-CAL–printed beams was substantially higher than that for SLA-printed beams (Fig. 4B), showing that the fracture strengths of CAL-printed components are more tightly distributed (*27*).

To demonstrate the mechanical properties of a more complex micro-CAL–printed object, we fabricated a Howe truss (*28*), subjected it to three-point bend loading, and found that it achieved a fracture stress of 187.7 MPa (Fig. 4C and fig. S10). VAM limits the creation of microcracks and indentations that would otherwise compromise fracture strength. The mechanical characterization shows that micro-CAL can produce complex, high-strength fused silica components with superior reliability to other AM modalities. In the future, micro-CAL could be used to investigate new high-strength lattices that exploit silica's high intrinsic strength and strain at failure in the absence of large flaws (*29*).

Fused silica glass microfluidic devices offer many advantages over polymeric devices, including high resistance to temperature and harsh acids and organic solvents as well as high optical transmission over an extended UV, visible, and infrared range. However, conventional fabrication techniques such as planar lithographic processes require toxic fluoric etchants and are largely limited to two dimensions (*30*). With micro-CAL, we achieved rapid free-form fabrication of perfusable branched 3D microfluidics with low surface roughness, high transparency, and channel diameters and wall thicknesses as low as 150 and 85 μm, respectively (Fig. 4, D and E, and fig. S12). These properties show that micro-CAL has the potential to advance the fabrication

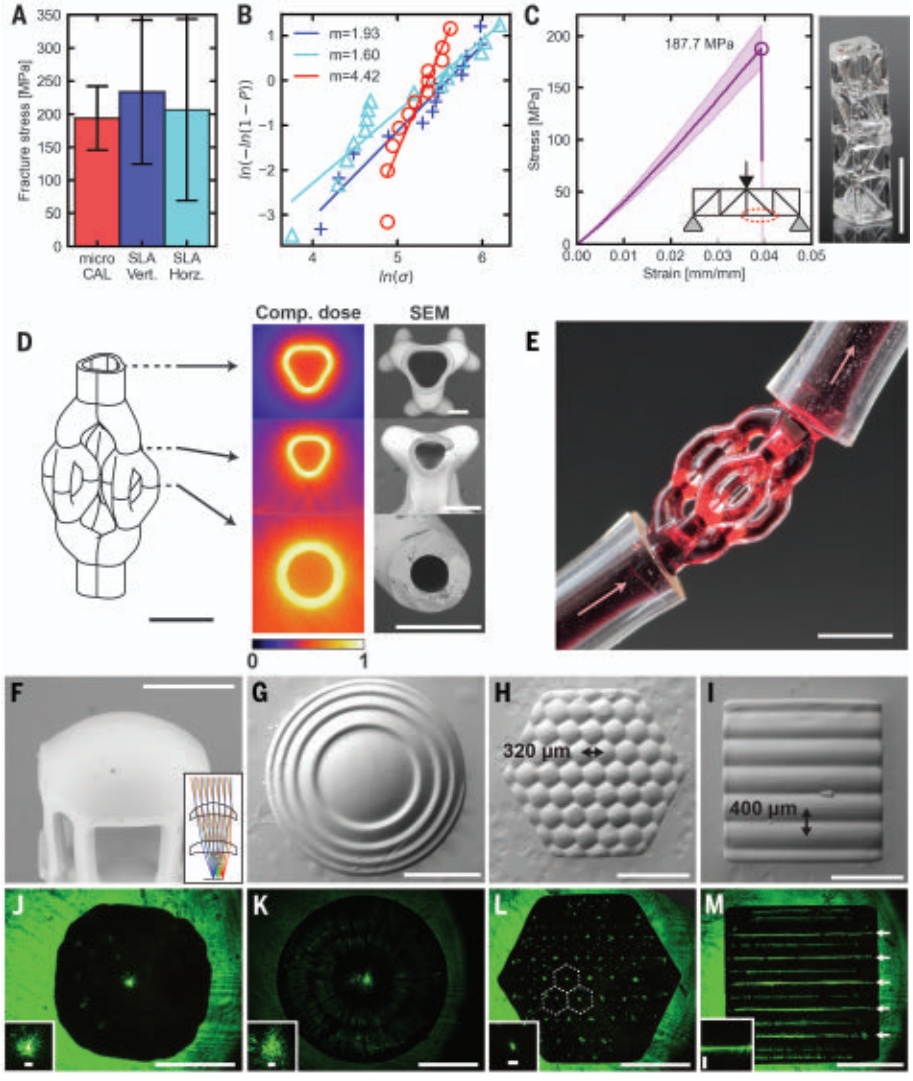


Fig. 4. Applications of glass CAL-printed microstructures. (**A**) Fracture stress of beams under flexure printed by micro-CAL and SLA in two different orientations: the long axis of the beam oriented vertically and horizontally with respect to the build plate. Data are means ± SD. *N* = 12, 14, and 16 for CAL, SLA vertical, and SLA horizontal, respectively. (**B**) Weibull modulus of the three types of printed beams, where *P* is the probability of failure and σ is the stress at failure. Color indicates print type and corresponds to the bar chart in (A). Each data point represents the fracture stress of an individual sample. (**C**) Three-point-bend test loading results (left) of a truss structure (right), where stress is the tensile stress in the bottom member as indicated by the red dashed oval in the schematic. The shaded region represents the bounded range of possible stresses due to variation in diameter of members. Scale bar is 2 mm. (**D**) Schematic of a trifurcated channel with normalized computational dose profiles (left) and SEM cross sections (right) at representative slices along a channel, revealing three levels of pore sizes: 750, 350, and 215 μm (top to bottom). Scale bars are 2 mm (schematic) and 500 μm (SEM micrographs). The color scale represents the normalized computational light dose. (**E**) Dyed liquid passed through the model demonstrates perfusability. Scale bar is 2 mm. (**F** to **I**) SEM micrographs of printed optical elements. (**J** to **M**) Point spread functions (PSFs) of the optical elements in (F) to (I) after focusing of 532-nm laser illumination. Insets show zoomed PSFs. Scale bars are 1 mm [(F) to (M)] and 50 μm [insets of (J) to (M)].

of microreactors, which are important for parallel drug screening and highly controlled flow synthesis (*3*).

The demand for more compact, lightweight, and high-quality cameras in consumer electronics and biomedical imaging pushes the development of advanced millimeter-scale optical systems. AM has enabled production of free-form refractive microlenses designed for specific applications, for example, foveated imaging. However, imaging elements made by layer-based techniques require post-processing such as polishing or coating to suppress the scattering induced by layer artifacts (*31*). With CAL, the 3D light dose possesses a radially and axially oriented gradient that has the effect of smoothening optical surfaces. We measured arithmetic mean line roughness (R_a) as low as 6 nm on as-fabricated lens surfaces using laser scanning confocal microscopy (fig. S17). We demonstrated printing of several refractive optical elements, including an air-spaced doublet aspheric lens optimized for operation at 532-nm wavelength, hexagonal and lenticular microlens arrays, and a spherical Fresnel lens (Fig. 4, F to I). The full-width half-maximum of the point spread function under collimated 532-nm illumination is less than 50 μm for each element. A low figure error in the range of 1 to 10 μm was achieved for spherical surfaces; however, a figure error of up to 60 μm persists for the aspheric lens surface, and imaging remains a challenge (figs. S15 and S16). These results suggest that although roughness is on par with commercial optics, figure error should be improved, perhaps by using in situ feedback and correction algorithms during printing (*17*, *32*).

The micro-CAL system we have developed enables manufacturing of structures with minimum feature sizes of 20 μm in polymer and 50 μm in fused silica, with excellent geometric freedom, low surface roughness, and high fracture strength and optical transparency in fused silica. Through optical engineering and specialized photopolymer development, we have established a glass fabrication framework that merges the facile processing of silica nanocomposites with layerless VAM and that could advance research in and industrial application of mechanical metamaterials, 3D microfluidics, and free-form optics.

REFERENCES AND NOTES

1. O. Solgaard, *Photonic Microsystems* (Springer, 2009).
2. B. C. Kress, I. Chatterjee, *Nanophotonics* **10**, 41–74 (2021).
3. F. Kotz, P. Risch, D. Helmer, B. E. Rapp, *Adv. Mater.* **31**, e1805982 (2019).
4. J. Luo *et al.*, *J. Manuf. Sci. Eng.* **139**, 061006 (2017).
5. C. Inamura, M. Stern, D. Lizardo, P. Houk, N. Oxman, *3D Print. Addit. Manuf.* **5**, 269–283 (2018).
6. K. C. Datsiou *et al.*, *J. Am. Ceram. Soc.* **102**, 4410–4414 (2019).
7. J. Lei, Y. Hong, Q. Zhang, F. Peng, H. Xiao, in *2019 Conference On Lasers Electro-Optics (CLEO)* (IEEE, 2019).
8. K. Sasan *et al.*, *ACS Appl. Mater. Interfaces* **12**, 6736–6741 (2020).
9. F. Kotz *et al.*, *Nature* **544**, 337–339 (2017).
10. I. Cooperstein, E. Shukrun, O. Press, A. Kamyshny, S. Magdassi, *ACS Appl. Mater. Interfaces* **10**, 18879–18885 (2018).
11. D. G. Moore, L. Barbera, K. Masania, A. R. Studart, *Nat. Mater.* **19**, 212–217 (2020).
12. F. Kotz *et al.*, *Adv. Mater.* **33**, e2006341 (2021).
13. J. Klein *et al.*, *3D Print. Addit. Manuf.* **2**, 92–105 (2015).
14. D. Yang, L. Liu, Q. Gong, Y. Li, *Macromol. Rapid Commun.* **40**, e1900041 (2019).
15. M. Shusteff *et al.*, *Sci. Adv.* **3**, eaao5496 (2017).
16. B. E. Kelly *et al.*, *Science* **363**, 1075–1079 (2019).
17. D. Loterie, P. Delrot, C. Moser, *Nat. Commun.* **11**, 852 (2020).
18. I. Bhattacharya, J. Toombs, H. Taylor, *Addit. Manuf.* **47**, 102299 (2021).
19. C. M. Rackson *et al.*, *Addit. Manuf.* **48**, 102367 (2021).
20. Materials and methods are available as supplementary materials.
21. E. T. Denisov, I. B. Afanas'ev, *Oxidation and Antioxidants in Organic Chemistry and Biology* (CRC Press, 2005).
22. C. C. Cook *et al.*, *Adv. Mater.* **32**, e2003376 (2020).
23. C. Maibohm *et al.*, *Sci. Rep.* **10**, 8740 (2020).
24. M. Thiel, M. S. Rill, G. von Freymann, M. Wegener, *Adv. Mater.* **21**, 4680–4682 (2009).
25. N. Sreedhar *et al.*, *Desalination* **425**, 12–21 (2018).
26. X. Zheng *et al.*, *Science* **344**, 1373–1377 (2014).

27. W. Weibull, *A Statistical Theory of the Strength of Materials* (Generalstabens Litografiska Anstalts Förlag, Stockholm, 1939).
28. W. Howe, Truss frame for bridges, US patent 1685 (1840).
29. C. R. Kurkjian, P. K. Gupta, R. K. Brow, *Int. J. Appl. Glass Sci.* **1**, 27–37 (2010).
30. C. Iliescu, H. Taylor, M. Avram, J. Miao, S. Franssila, *Biomicrofluidics* **6**, 016505 (2012).
31. X. Chen *et al.*, *Adv. Mater.* **30**, e1705683 (2018).
32. C. C. Li, J. Toombs, H. Taylor, in *Proceedings of the Symposium on Computational Fabrication 2020 (SCF '20)* (Association for Computation Machinery, 2020).

ACKNOWLEDGMENTS

We thank the staff at the University of California (UC), Berkeley, Electron Microscopy Lab for advice in electron microscopy sample preparation and data collection. We thank P. Risch and M. Sanjaya for providing reference samples. We thank A. Warmbold for the thermogravimetric analysis measurement. We thank S. Kumar's lab at UC Berkeley for allowing use of the rheometer. We would like to thank the anonymous reviewers for their time and valuable critique.

Funding: This work was funded by the National Science Foundation under cooperative agreement no. EEC-1160494 (J.T.T., C.C.L., H.K.T.); the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (grant agreement nos. 816006 to M.L. and 825521 to S.J.); the Carl Zeiss Foundation as a part of the Research Cluster "Interactive and Programmable Materials (IPROM)"; the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) through the Centre for Excellence livMatS Exec 2193/1 – project number 390951807 (B.E.R.); the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) project number 455798326 (F.K.-H.); and the Lawrence Livermore National Laboratory Directed Research and Development program. The work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344 (LLNL-JRNL-826682) (C.C.C.).
Author contributions: Conceptualization: J.T.T., F.K.-H., H.K.T.; Methodology: J.T.T., M.L., F.K.-H.; Investigation: J.T.T., M.L., C.C.C., S.J., C.C.L.; Visualization: J.T.T.; Funding acquisition: H.K.T., F.K.-H., B.E.R.; Project administration: H.K.T., F.K.-H., J.T.T.; Supervision: F.K.-H., H.K.T., B.E.R.; Writing – original draft: J.T.T.; Writing – review and editing: All authors. **Competing interests:** H.K.T. holds US patent 10,647,061 relating to computed axial lithography. B.E.R. and F.K.-H. hold US patent 10,954,155 B2 relating to the silica nanocomposite material described in this paper. F.K.-H. and B.E.R. are co-founders of, and have an equity interest in, Glassomer GmbH. The authors declare that they have no other competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Code used to analyze and visualize data is available upon request.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm6459
Materials and Methods
Supplementary Text
Figs. S1 to S23
Tables S1 to S9
References (*33*–*46*)

30 September 2021; accepted 23 February 2022
10.1126/science.abm6459

MATERIALS SCIENCE

Complex morphologies of biogenic crystals emerge from anisotropic growth of symmetry-related facets

Emanuel M. Avrahami[1], Lothar Houben[2], Lior Aram[1], Assaf Gal[1]*

Directing crystal growth into complex morphologies is challenging, as crystals tend to adopt thermodynamically stable morphologies. However, many organisms form crystals with intricate morphologies, as exemplified by coccoliths, microscopic calcite crystal arrays produced by unicellular algae. The complex morphologies of the coccolith crystals were hypothesized to materialize from numerous crystallographic facets, stabilized by fine-tuned interactions between organic molecules and the growing crystals. Using electron tomography, we examined multiple stages of coccolith development in three dimensions. We found that the crystals express only one set of symmetry-related crystallographic facets, which grow differentially to yield highly anisotropic shapes. Morphological chirality arises from positioning the crystals along specific edges of these same facets. Our findings suggest that growth rate manipulations are sufficient to yield complex crystalline morphologies.

Control over nanoscale morphologies of crystalline materials is connected to their physical properties and thus their potential applications (*1*–*3*). However, the inherent thermodynamic properties of the crystalline lattice dictate a strong tendency toward specific low-energy facets, resulting in characteristic shapes (habits) (*4*). In contrast, many organisms evolved the ability to form intricate crystalline structures with hierarchical organization, out of very simple materials and under ambient conditions. In such biomineralization processes, the polymorph of the crystal, its nucleation site, orientation, and eventual morphology, are all under strict control (*5*–*7*).

Coccoliths—micrometer-sized calcite (calcium carbonate) scales formed by unicellular algae called coccolithophores—are a prime example of biological control over crystal morphogenesis. Each coccolith is composed of crystalline subunits and has intricate species-specific morphology (*8*, *9*). Coccoliths are made intracellularly within a specialized vesicle known as the coccolith vesicle, into which calcium and carbonate ions are delivered (*10*, *11*). Inside the coccolith vesicle, crystals nucleate and grow around the rim of an organic base plate (*8*, *12*).

A common feature of coccolith architectures is an alternating arrangement of crystal units, as identified in the V/R model (*13*). According to the model, two unit types make up a coccolith—a V-unit and an R-unit—possessing either a vertical or a radial orientation of the calcite *c* axis relative to the base plate. The units initially possess a pseudo-rhombohedral morphology that closely resembles the thermodynamically stable {104} calcite rhombohedron (*14*, *15*). Nonetheless, upon completion, their morphology is highly convoluted, displaying various surfaces that deviate markedly from the simple rhombohedral habit (*8*, *14*).

The consensus view on coccolith morphogenesis is reliant on biomolecules as the "sculptor's toolkit." The rationale is that specific stereochemical interactions with the growing crystals allow such biomolecules to funnel the process away from the stable thermodynamic path, and into local kinetic minima, giving rise to potentially unlimited morphologies (*6*, *16*–*21*). Presumably, crystal nucleation is the result of epitaxy from the base plate, and crystal growth produces various types of crystallographic facets stabilized by "tailored" biomolecules. It has also been suggested that stereospecific interactions with chiral organic modifiers induce chirality to the habit of calcite (*21*–*23*).

To elucidate the morphological development of coccolith crystals, we investigated the large coccoliths of *Calcidiscus leptoporus*, which have a characteristic double-shield ultrastructure (Fig. 1A). To create a timeline of coccolith growth, we established a procedure for extracting intracellular coccoliths (ICCs). First, extracellular coccoliths of actively calcifying cells were removed with a short acid exposure. Next, a hypotonic solution was used to burst the cells, thus releasing their ICCs. By tuning the pH and chemistry of the hypotonic solution, we ensured that crystal morphologies were unaffected (figs. S1 to S3 and supplementary text). Therefore, the ICCs serve as "snapshots in time" of the dynamic development of the crystals.

Scanning electron microscopy (SEM) images of the extracted ICCs (Fig. 1) show a sequence of evolving intermediate morphologies from small, 100- to 200-nm rhombohedra to fully formed chiral coccoliths (see also fig. S4). The overall chirality of the structure is evident even in the arrangement of the initial units, which resembles the isotropic rhombohedral habit of calcite (Fig. 1, E and I). Two distinct surface types of the crystals were observed: (i) flat facets with straight edges, characterizing the distal sides of both shields (Fig. 1, purple arrowheads), and (ii) curved and smooth surfaces, characterizing the proximal sides of the two shields and stem region (Fig. 1, green arrowheads).

We used high-resolution electron tomography to investigate crystal morphologies of

[1]Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel. [2]Department of Chemical Research Support, Weizmann Institute of Science, Rehovot, Israel.
*Corresponding author. Email: assaf.gal@weizmann.ac.il

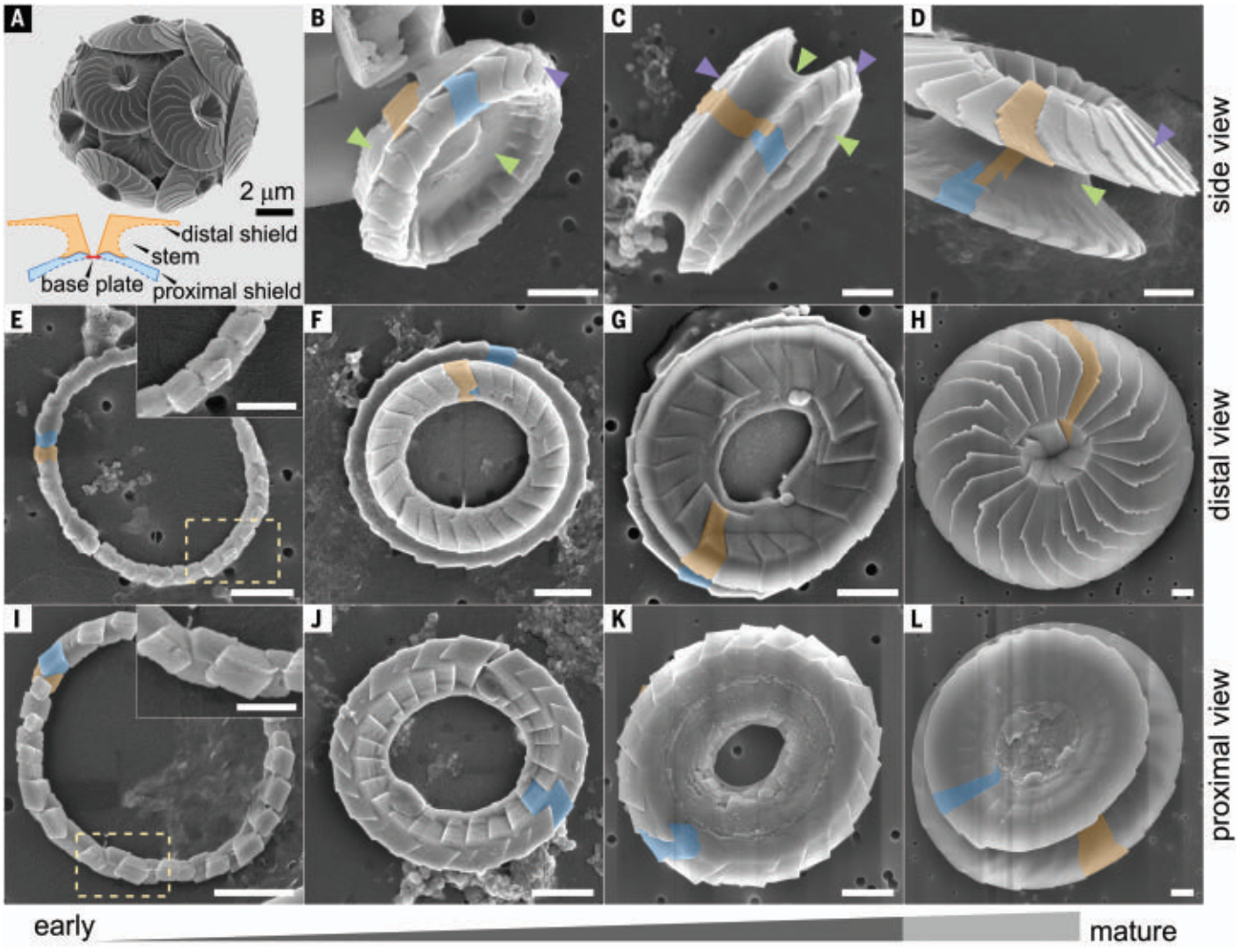


Fig. 1. Overview of crystal morphogenesis in developing coccoliths. (**A**) A *C. leptoporus* cell with a complete coccolith shell (top) and a schematic cross section of a mature coccolith (bottom). Here and in the remaining figures, the V-units, forming the distal shield, are in orange, and the R-units, forming the proximal shield, are in blue. Dashed lines indicate curved surfaces, continuous lines indicate flat surfaces, and a red line indicates the putative position of the base plate (*13*). (**B** to **L**) ICCs at four growth stages. Three different views highlight the architecture at each stage. Insets in (E) and (I) show a magnification of the dashed areas. Arrowheads indicate flat (purple) and curved (green) surfaces. The dark section in the schematic timeline at the bottom represents the morphological spectrum of ICCs analyzed in the following figures. Scale bars, 500 nm (insets, 200 nm).

both units in three dimensions and at different developmental stages. Tomograms of ICCs at different growth stages were collected in a scanning transmission electron microscope (STEM), using the high-angle annular dark-field (HAADF) detector, and reconstructed into 3D volumes (figs. S5 to S7) (*24*, *25*). 3D analysis of a coccolith at an early developmental stage showed that all crystal units expose flat crystallographic facets (Fig. 2). The dihedral angles between these surfaces and the angles between their edges are in agreement with those known for the {104} calcite rhombohedron (*26*), which suggests that only these stable crystallographic facets are expressed (fig. S8).

We observed that R-units are situated on their acute edges, aligned along the circumference of the coccolith ring (Fig. 2A)—an arrangement that is in agreement with observations in other species (*15*, *27*). This is interesting for two reasons: (i) As a result of geometrical considerations, and in contrast to the conventional V/R model, aligning such {104} rhombohedra on their acute edges enforces a subradial orientation of the crystals' *c* axes, breaking radial symmetry and conveying chirality to the emergent structure (Fig. 2A, cyan arrow); (ii) it challenges the concept of epitaxy, as the crystals should have a facet, and not an edge, parallel to the nucleating surface (i.e., the base plate). Although less clear in the V-units (initial crystals appear less rhombohedral), we also see that the crystals possess a subvertical tilt of their *c* axes—a result of orienting the rhombohedra on their obtuse edge (Fig. 2B and fig. S9). Given that our data lack information on the mechanisms that lead to this regulated crystal orientation, the role of the base plate as a nucleating surface remains an open question.

To relate the morphological information to the crystallography of the crystals, we analyzed adjacent R-units from an ICC with annular dark-field (ADF) STEM coupled with scanning nanobeam electron diffraction (NBED) (*28*), which collects a diffraction pattern from every point the beam raster traverses. The analysis confirmed a relative tilt between the units, as well as the subradial deviation of each *c* axis relative to the coccolith circumference (Fig. 2, C to E). These analyses of early-stage ICCs allow us to refine the "classical" V/R model, which centers on *c*-axis directions and is achiral, to a more accurate crystallographic representation of crystal

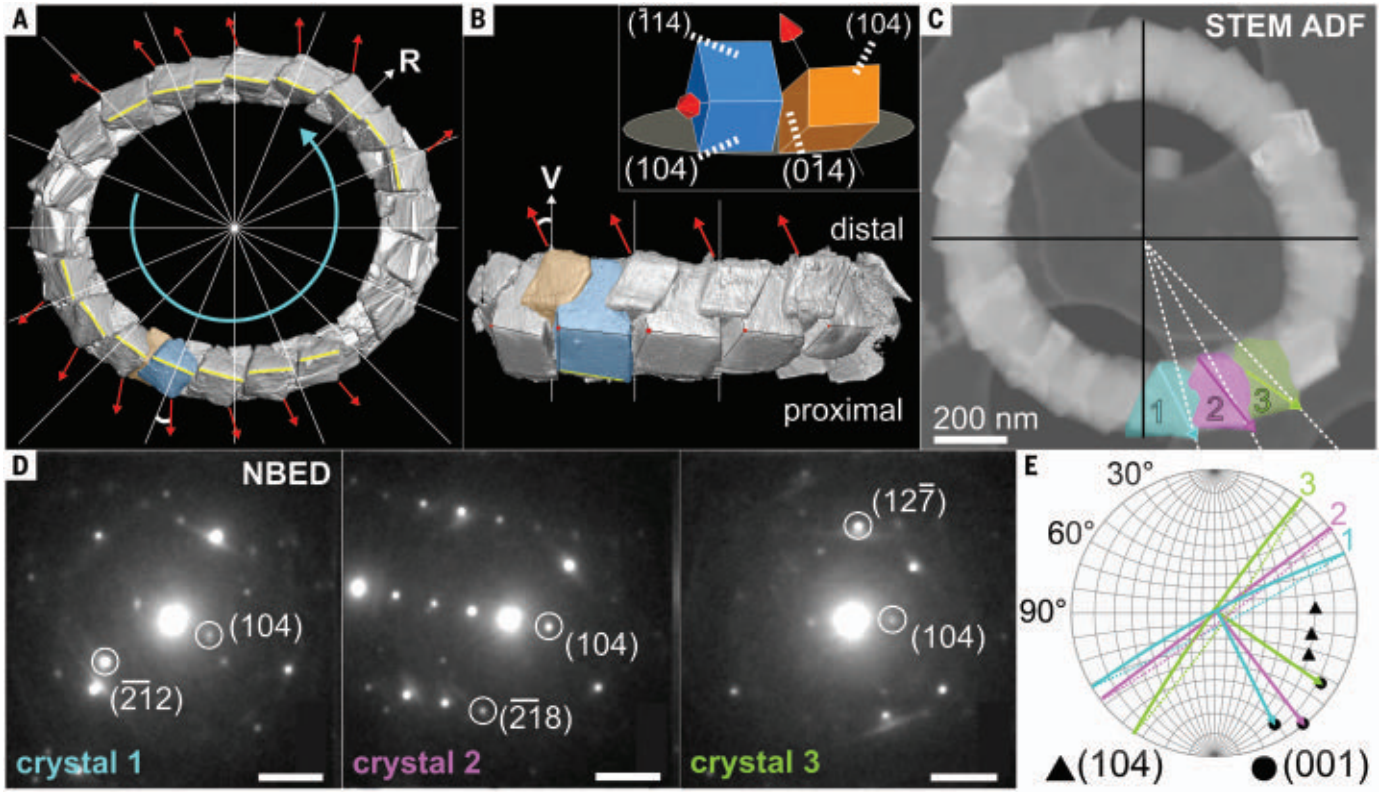


Fig. 2. Alignment of crystal edges at the coccolith circumference results in coccolith chirality. (**A**) 3D rendering of an early-stage ICC, viewed from the proximal side. Yellow lines outline the acute edges on which the R-units are situated along the circumference of the base plate. The *c* axes of the R-units (red arrows) and their subradial tilts (note deviations from the lines illustrating radial directions) are shown, as well as the emerging ultrastructural chirality (cyan arrow). (**B**) Side view of the coccolith in (A), showing the subvertical tilt in *c* axes of the V-units (angles between the vertical lines and red arrows). Inset shows schematic {104} rhombohedra with crystallographic annotations, situated on obtuse (R-unit) or acute (V-unit) edges. (**C**) STEM ADF image of an ICC at a stage similar to that in (A); colored crystals were analyzed by NBED, and the derived *c*-axis orientations [in (E)] are indicated as colored arrows. (**D**) NBED patterns associated with the crystal units marked in (C). Diffraction scale bars, 5 nm^{-1}. (**E**) The relative orientation of the three R-units in (C) is shown in a stereographic projection of the (104) and (001) poles to the coccolith plane.

rhombohedra resting on their acute/obtuse edges. This view consolidates two crystallographic features into one basic architecture, in which both the tilted *c* axes and the chiral ultrastructure originate from the initial positioning of the crystals.

To understand the ways by which coccolith crystals grow and interlock, we analyzed the morphology of individual crystal units in detail. Five coccoliths, reflecting the stages along coccolith growth that are suitable for tomography (see Fig. 1, bottom bar), were partially segmented (Fig. 3 and movies S1 to S5). The derived "timeline" revealed several key aspects: (i) Both unit types demonstrate a transition from relatively isotropic rhombohedra to mature anisotropic crystals (Fig. 3, A and B); (ii) both unit types feature facets that appear crystallographic throughout their growth, while some areas—the stem region, the proximal sides of the shields, and the interfaces between neighboring crystals—maintain a curved morphology (Fig. 3C); (iii) the dihedral angles between crystallographic-like facets throughout crystal growth all correspond to the {104} habit (fig. S8).

This timeline demonstrates how the equivalent {104} facets of initial crystals develop in an anisotropic manner, giving rise to mature {104} facets with very different sizes (Fig. 3, D and E). These observations show that the complex morphology of the crystals is not the result of various types of crystallographic planes, but rather of differential growth of the chemically equivalent {104} facets.

The observations that the crystals are growing with the expression of {104} facets exclusively (Fig. 3), and that these {104} facets grow at different speeds, raise a critical question regarding the factors responsible for this symmetry breaking. This conundrum arises from the symmetry and chemical equivalence of all six {104} facets, such that no single facet possesses an inherent growth rate that differs from the others (i.e., calcium and carbonate ions should display no association or dissociation bias toward any specific {104} facet) (*26*, *29*).

To understand how the anisotropy of these chemically equivalent facets emerges, we analyzed growth patterns of specific facets. Two distinct patterns were observed: (i) differential growth of symmetry-related facet pairs of an individual crystal unit [e.g., ($\bar{1}$14) and (104) (Fig. 4A and fig. S10)], where one facet grows faster than its opposite and/or adjacent facets, resulting in an anisotropic motif; and (ii) differential growth of facets from two different unit types (R and V) facing the same environment (Fig. 4B). In the latter case, facets first appear level with one another, yet end up with the V-units repeatedly outgrowing the R-units (Fig. 4B, compare insets). Both examples show two chemically equivalent facets that, for some reason, differ in their growth rates.

Within a homogeneous solution, anisotropic growth of equivalent crystallographic facets is incompatible with their identical growth kinetics. However, on the atomic scale, calcite growth proceeds via both acute and obtuse steps, each having different growth kinetics (*29*–*31*). Therefore, nanoscale inhomogeneity in the environment in which the crystals grow can result in growth anisotropy. It was shown in several coccolithophore species that crystallization occurs at extreme confinement, where only tens of nanometers separate the crystals

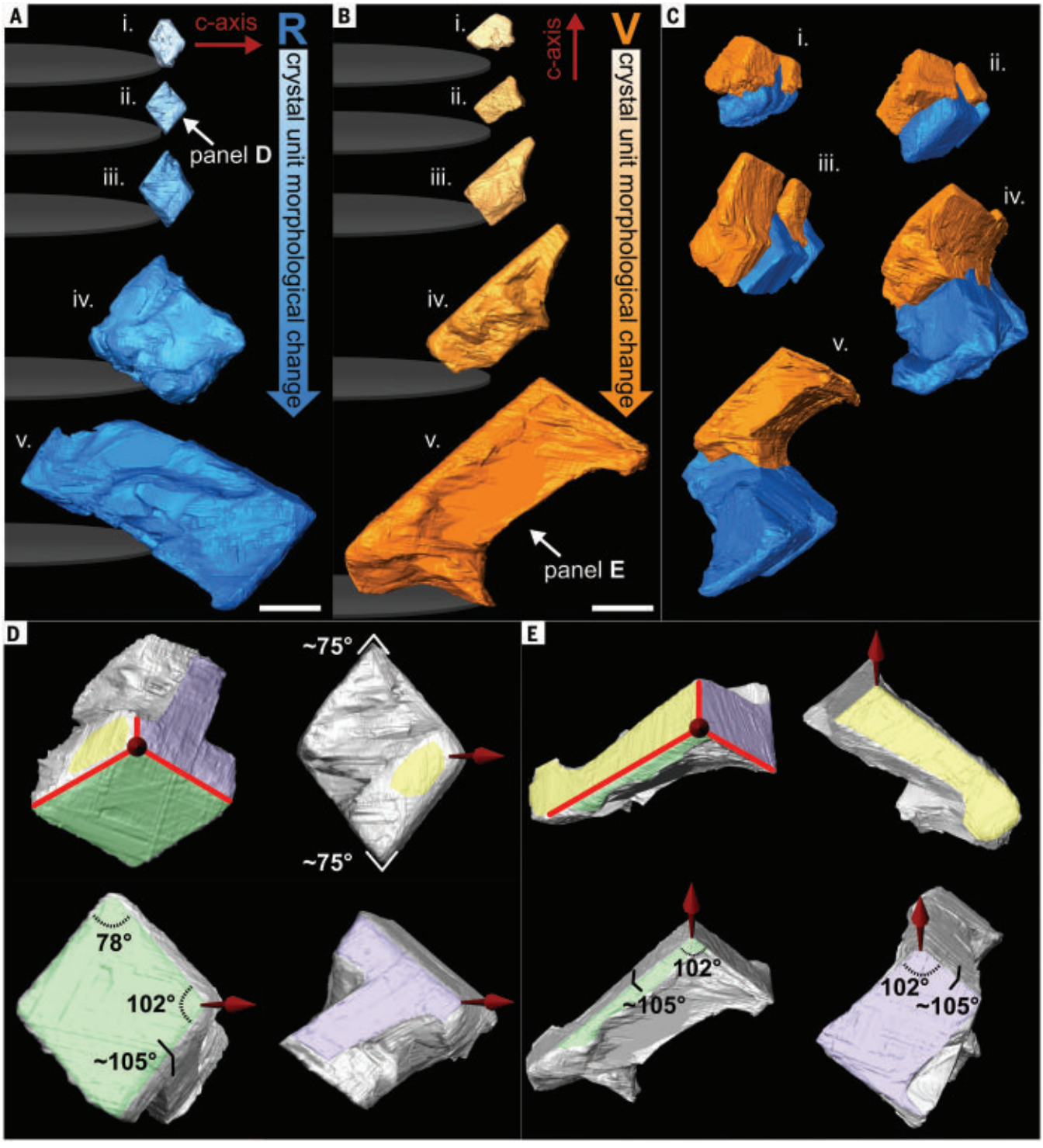


Fig. 3. Morphological "timeline" of coccolith crystal growth shows evolution of anisotropy with conserved {104} facets. (**A** and **B**) 3D volume rendering of crystals from five developmental stages indicated as i. to v. (color gradient denotes developmental sequence). The units are viewed tangentially to the coccolith ring to emphasize the {104} habit, and they show the transition from isotropic to anisotropic morphologies. Gray discs represent the schematic location of the base plate. Scale bars, 100 nm. (**C**) Four interlocked crystal units from the same data sets. The crystals are viewed from outside the coccolith perimeter, and sizes are not to scale. (**D** and **E**) An early-stage R-unit (D) and a mature V-unit (E) at four different views: down their *c* axes (upper left of both panels), and face-on to each of the facets joining at the *c*-axis apex. These perspectives highlight the initial isotropic habit, mostly disturbed by the interlocking of the crystals, which transforms to an anisotropic habit that is still dominated by the same facets. Dihedral angles between facets (continuous lines) and angles between edges (dashed lines) are indicated. Red arrows denote directions of *c* axes.

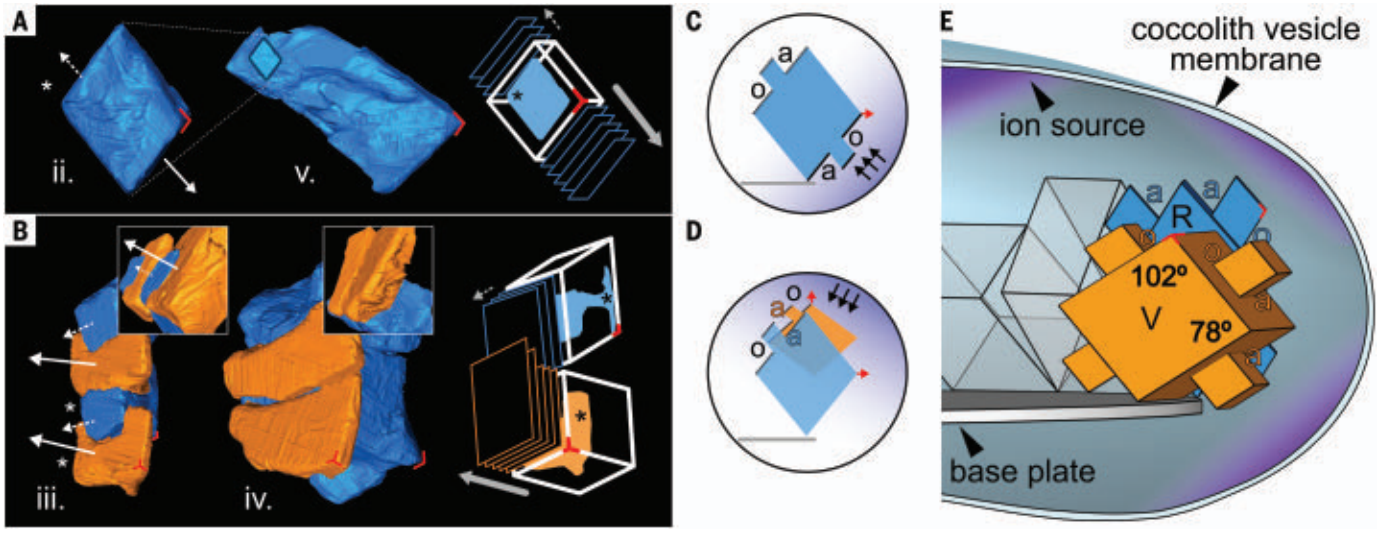


Fig. 4. Development of crystal anisotropy due to differential growth rates of equivalent facets. (**A** and **B**) Two sets of crystals at different growth stages (same as in Fig. 3). In (A), the former is schematically superimposed on the latter. In addition, a schematic {104} rhombohedron (in white) that is superimposed over a silhouette of the early unit (asterisks) shows the differentially growing facets. Continuous arrows indicate fast-growing facets; dashed arrows designate slow-growing facets. (**C** and **D**) Illustrations of the possible conditions inside the cell leading to the observed growth regimes in (A) and (B). A putative gradient in ion concentration is shown in purple and with black arrows. In scenario (C) the directional flux leads to accelerated growth of one facet toward it, unlike a different facet on the same crystal that is farther from the ion source. In scenario (D), two different units have different step orientations facing a similar gradient. In this case, their differential growth rate will determine which facet grows faster. a, acute step; o, obtuse step. (**E**) A model showing local ion concentration gradients (purple) in the confined environment of the coccolith vesicle. This anisotropic environment enables different atomic steps of adjacent crystals to experience different solution conditions. In all panels, red arrows and lines at crystal apexes indicate *c*-axis directions.

and the vesicle membrane (*8*, *27*, *32*). Such confinement manifests directly in crystal morphology as noncrystallographic surfaces that result from a physical block of growth by the delimiting membrane of the coccolith vesicle (*27*). We propose that confinement also influences crystal growth indirectly by creating a graded nanoenvironment within the coccolith vesicle. For example, a concentration gradient may arise from localized ion fluxes generated by ion transporters on the vesicle membrane (*33*). It remains imperative to characterize, chemically and structurally, the cellular environment and its interactions with the growing crystals.

Figure 4, C to E, illustrates how such a concentration gradient differentially affects growth steps at the atomic scale, leading to different growth kinetics of equivalent facets that result in an anisotropic growth. For example, when one facet of a crystal experiences a higher ion concentration than another, it will grow faster toward the ion source (Fig. 4C). Even more interesting is when two neighboring facets of different crystals present different geometries of their atomic steps toward the ion gradient (Fig. 4D), resulting in faster growth of one of the crystals. The differences between step orientations in the presence of a nanoscale gradient (Fig. 4E) break the symmetry between the adjacent crystals and can explain their anisotropic growth.

Coccolith crystal growth is not a process that stems from multiple manipulations of crystallographic growth; rather, it hinges on the various consequences that emerge from the stable habit of calcite and its rhombohedral geometry. Such a growth regime can be controlled by the rate and location of ion transport, rather than by "tailored" modifications for specific crystal facets. One can envision how alterations in the initial conditions of coccolith assembly (e.g., unit orientation, unit spacing, ion flux direction, or membrane position during growth) can markedly affect the final coccolith morphology.

REFERENCES AND NOTES

1. J. Aizenberg, A. Tkachenko, S. Weiner, L. Addadi, G. Hendler, *Nature* **412**, 819–822 (2001).
2. F. Nudelman, N. A. J. M. Sommerdijk, *Angew. Chem. Int. Ed.* **51**, 6582–6596 (2012).
3. D. Palin *et al.*, *J. Am. Chem. Soc.* **143**, 3439–3447 (2021).
4. J. J. De Yoreo, P. J. Vekilov, *Rev. Mineral. Geochem.* **54**, 57–93 (2003).
5. B. Bayerlein *et al.*, *Nat. Mater.* **13**, 1102–1107 (2014).
6. S. Weiner, L. Addadi, *Annu. Rev. Mater. Res.* **41**, 21–40 (2011).
7. W. Jiang *et al.*, *Science* **368**, 642–648 (2020).
8. J. R. Young, *Rev. Mineral. Geochem.* **54**, 189–215 (2003).
9. F. M. Monteiro *et al.*, *Sci. Adv.* **2**, e1501822 (2016).
10. D. E. Outka, D. C. Williams, *J. Protozool.* **18**, 285–297 (1971).
11. M. E. Marsh, *Comp. Biochem. Physiol. B* **136**, 743–754 (2003).
12. A. Gal *et al.*, *Science* **353**, 590–593 (2016).
13. J. R. Young, J. M. Didymus, P. R. Brown, B. Prins, S. Mann, *Nature* **356**, 516–518 (1992).
14. J. M. Didymus, J. R. Young, S. Mann, *Proc. R. Soc. B* **258**, 237–245 (1994).
15. J. M. Walker, B. Marzec, N. Ozaki, D. Clare, F. Nudelman, *J. Struct. Biol.* **210**, 107476 (2020).
16. L. Addadi *et al.*, *Nature* **296**, 21–26 (1982).
17. A. M. Belcher *et al.*, *Nature* **381**, 56–58 (1996).
18. D. B. DeOliveira, R. A. Laursen, *J. Am. Chem. Soc.* **119**, 10627–10631 (1997).
19. F. C. Meldrum, *Int. Mater. Rev.* **48**, 187–224 (2003).
20. J. J. De Yoreo *et al.*, *Science* **349**, aaa6760 (2015).
21. W. Jiang, M. S. Pacella, H. Vali, J. J. Gray, M. D. McKee, *Sci. Adv.* **4**, eaas9819 (2018).
22. S. Mann, N. Sparks, *Proc. R. Soc. London Ser. B* **234**, 441–453 (1988).
23. K. Henriksen, S. L. S. Stipp, J. R. Young, M. E. Marsh, *Am. Mineral.* **89**, 1709–1716 (2004).
24. J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, *J. Struct. Biol.* **116**, 71–76 (1996).
25. D. N. Mastronarde, *J. Struct. Biol.* **152**, 36–51 (2005).
26. R. J. Reeder, *Rev. Mineral. Geochem.* **11**, 1–47 (1983).
27. Y. Kadan, F. Tollervey, N. Varsano, J. Mahamid, A. Gal, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2025670118 (2021).
28. M. W. Tate *et al.*, *Microsc. Microanal.* **22**, 237–249 (2016).
29. C. A. Orme *et al.*, *Nature* **411**, 775–779 (2001).
30. E. Ruiz-Agudo, C. V. Putnis, L. Wang, A. Putnis, *Geochim. Cosmochim. Acta* **75**, 3803–3814 (2011).
31. M. De La Pierre, P. Raiteri, A. G. Stack, J. D. Gale, *Angew. Chem. Int. Ed.* **56**, 8464–8467 (2017).
32. A. R. Taylor, M. A. Russell, G. M. Harper, T. F. T. Collins, C. Brownlee, *Eur. J. Phycol.* **42**, 125–136 (2007).
33. A. R. Taylor, C. Brownlee, G. Wheeler, *Annu. Rev. Mar. Sci.* **9**, 283–310 (2017).
34. E. M. Avrahami, L. Houben, L. Aram, A. Gal, Complex morphologies of biogenic crystals emerge from anisotropic growth of symmetry-related facets. Dryad (2022); doi:10.5061/dryad.zcrjdfndp.

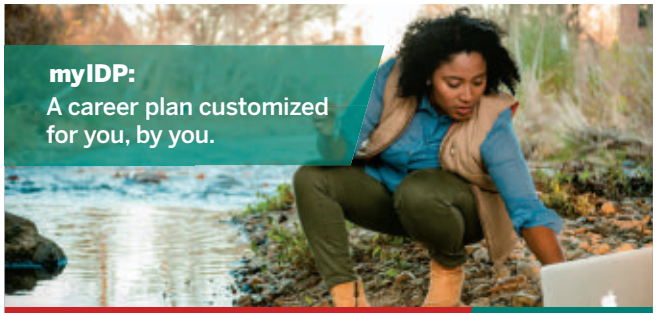
ACKNOWLEDGMENTS

We thank E. Shimoni for dedicated assistance with HAADF-STEM acquisitions, L. Addadi for helpful discussions, and E. Komendacka for designing the schematic 3D model of a coccolith ring. **Funding:** Supported by Israel Science Foundation grant 697/19. **Author contributions:** E.M.A. performed coccolith extractions, HAADF-STEM data collection, computational reconstructions, segmentations and 3D visualizations; L.H. performed the NBED measurements; L.A. performed HAADF-STEM imaging at cryogenic conditions; E.M.A. analyzed the results with the guidance of A.G; A.G. supervised the research; A.G. and E.M.A. wrote the paper with the assistance of L.H. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Electron microscopy raw data, as well as volume renderings and segmentation files, are available through the Dryad digital repository (*34*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm1748
Materials and Methods
Supplementary Text
Figs. S1 to S11
Movies S1 to S5

31 August 2021; accepted 15 February 2022
10.1126/science.abm1748



For your career in science, there's only one Science

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

ScienceCareers In partnership with:

FASEB · University of Massachusetts Medical School · UCSF University of California San Francisco · UIC · Burroughs Wellcome Fund

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

ARL Distinguished Postdoctoral Fellowships

The Army Research Laboratory (ARL) Distinguished Postdoctoral Fellowships provide opportunities to pursue independent research in ARL laboratories. Fellows benefit by working alongside some of the nation's best scientists and engineers, while enhancing the mission and capabilities of the U.S. Army and the warfighter in times of both peace and war.

Fellows must display extraordinary abilities in scientific research and show clear promise of becoming future leaders. Candidates are expected to have already successfully tackled a major scientific or engineering problem or to have provided a new approach or insight, evidenced by a recognized impact in their field.

Fellowships are one-year appointments, renewable for up to three based on performance. The award includes a $115,000 annual stipend, health insurance, paid relocation, and a professional travel allowance. Applicants must have completed all requirements for a Ph.D. or Sc.D. degree by October 1, 2022 and may not be more than five years beyond their doctoral degree as of the application deadline. For more information and to apply visit **www.nas.edu/arl**.

Online applications must be submitted by June 15, 2022 at 5 PM Eastern Time.

U.S. ARMY · DEVCOM Army Research Laboratory

Find your next job at ScienceCareers.org

Whether you're looking to get ahead, get into, or just plain get advice about careers in science, there's no better or more trusted authority. Get the scoop, stay in the loop with *Science* Careers.

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

By **Mingde Zheng**

Listen to your body

After drawing blood and running a CT scan, the doctor had questions for me. "What is a typical day like for you?" he inquired. "I am a graduate student with a packed schedule that usually keeps me at work late into the night," I replied. Next, he turned to diet. I paused when this question came, embarrassed by my answer. "I only have coffee for breakfast. For lunch and dinner, I usually grab something from a fast-food vendor on campus." The doctor seemed aghast. As more questions followed about my stress levels and lifestyle, my unhealthy state began to sink in.

It was easier to live a well-rounded life and stay healthy when I was an undergraduate student. I had fewer responsibilities—succeeding in classes was the main priority. I lived in a dorm with resident assistants who advised us about our personal lives, showing us where to eat and exercise. The cafeteria served healthy food options. And I had a built-in group of friends through the dorm, which made it easy to participate in social activities.

That all changed in graduate school. I was laser focused on my dissertation project and doing what I could to become a successful scientist. I lived off campus, and I didn't feel I had the time or energy to shop for groceries and cook. I stopped exercising and didn't spend as much time socializing. My drive to succeed blinded me from seeing any potential problems with my lifestyle.

"I am thankful I had an early warning that I was neglecting my physical health."

Sometimes, I could feel that I was becoming sluggish. But instead of giving my mind and body a rest, I drank energy boosters to maintain my stamina and continued to work long hours. I became an abusive overlord of my body. But it could only take so much before it crashed and burned.

As I progressed through my Ph.D. program, I gained an unhealthy amount of weight. I got tired easily. Energy drinks no longer worked, and my body would react violently to them, with nausea, chills, and shakes. I often caught colds or the flu. And I started to have sporadic abdominal pains, which I tried my best to ignore.

Eventually, the abdominal pains were so bad they drove me to the emergency room. That's when the doctor peppered me with questions. As my wife sat terrified by my side, he diagnosed me with a gastrointestinal disorder and told me in no uncertain terms that I would need to change my lifestyle. No pill would fix my problem; in the long term, eating well, exercising, and trying to minimize stress would be the only way to keep the pain from recurring and turning into something worse.

I worried that if I eased up on the gas pedal at work, I wouldn't be as productive. But after hearing the doctor's warning, I knew I had no choice. I stopped working late into the night, which gave me more time to relax, sleep, and prepare my own meals. I read up on dietary recommendations and began to choose nutritious and healthy foods. I also gave up energy drinks and switched to tea.

Many times, I was tempted to go back to my old routine. I placed sticky notes in my office and at home to remind myself that if I didn't make healthy choices, I'd suffer consequences. Over time, though, it became easier, especially once my habits became more ingrained and I began to notice positive changes in my life.

It's been 4 years now and, to my surprise, I have not only noticed gains in my health, I have also found it easier to be productive at work. I'm no longer plagued by stomach pains and constant fatigue. And I'm able to remain alert throughout the day without having to rely on energy boosters.

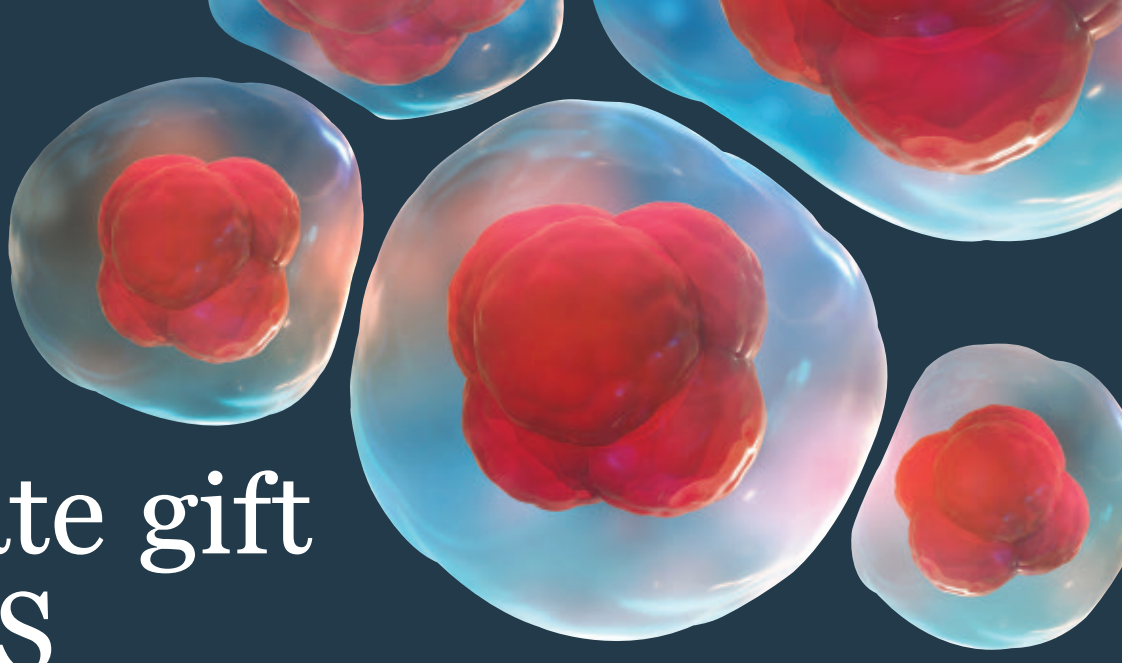
My unhealthy lifestyle may not have been the root cause of my health issues—I may have become ill regardless. But I am thankful I had an early warning that I was neglecting my physical health. I've noticed many other scientists making similar mistakes, and I worry their wake-up call won't come until later in life, when it will be harder to reverse the damage. Our work as scientists is important, but we can only do that work if we take care of our bodies. ■

Mingde Zheng is a research scientist at Nokia Bell Labs. Do you have an interesting career story to share? Send it to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

1848 SOCIETY

AAAS



An estate gift to AAAS

Going all the way back to 1848, our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

Today, we are dedicated to advocating for science and scientific evidence to be fully and positively integrated into public policy and for the community to speak with one voice to advance science and engineering in the United States and around the world.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you will become a member of our 1848 Society and will help fuel our work on behalf of science and society – including publishing the world's most promising, innovative research in the *Science* family of journals and engaging in the issues that matter locally, nationally and around the world.

"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."

—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science
Federal Tax ID Number: 53-0196568
Address: 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

cut here

Yes, I would like more information about joining the AAAS 1848 Society.

1848 SOCIETY

AAAS

PLEASE CONTACT ME AT:

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

RETURN THIS FORM TO:
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

AAAS.ORG/COMMUNITY

AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY
AAAS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE